

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

# OM protein - protein search, using sw model

Run on: June 2, 2005, 01:27:36 ; Search time 28.125 Seconds  
(without alignments)  
39.813 Million cell updates/sec

Title: US-10-764-235-2

Perfect score: 68

Sequence: 1 KKAADVLLPVLAAAP 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 750 summaries

Database : Issued Patents AA.\*

- 1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pdp.\*
- 2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pdp.\*
- 3: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pdp.\*
- 4: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pdp.\*
- 5: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.pdp.\*
- 6: /cgn2\_6/ptodata/1/iaa/backfiles1.pdp.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	68	100.0	15	4	US-09-671-089-2
2	68	100.0	16	4	US-09-671-089-3
3	68	100.0	16	4	US-09-671-089-48
4	68	100.0	19	4	US-09-671-089-4
5	64	94.1	16	4	US-09-671-089-5
6	54	79.4	12	3	US-09-186-170-1
7	54	79.4	12	4	US-09-562-868-1
8	54	79.4	12	4	US-09-997-465B-4
9	54	79.4	12	4	US-10-083-889-8
10	54	79.4	12	4	US-10-116-288A-1
11	54	79.4	12	4	US-09-671-089-1
12	54	79.4	12	4	US-09-671-089-14
13	54	79.4	12	4	US-10-144-549-19
14	54	79.4	12	4	US-10-116-288A-22
15	54	79.4	25	4	US-10-116-288A-20
16	53	77.9	14	4	US-09-671-089-7
17	53	77.9	14	4	US-09-671-089-58
18	51	75.0	17	4	US-09-671-089-24
19	50	73.5	11	3	US-09-186-170-9
20	50	73.5	11	4	US-09-562-868-9
21	50	73.5	11	4	US-10-116-288A-9
22	50	73.5	11	4	US-09-671-089-19
23	50	73.5	13	4	US-09-671-089-10
24	47	69.1	11	3	US-09-186-170-5
25	47	69.1	11	4	US-09-562-868-5
26	47	69.1	11	4	US-10-116-288A-5
27	47	69.1	11	4	US-09-671-089-15
28					Sequence 2, Appli
29					Sequence 3, Appli
30					Sequence 48, Appli
31					Sequence 4, Appli
32					Sequence 5, Appli
33					Sequence 1, Appli
34					Sequence 8, Appli
35					Sequence 1, Appli
36					Sequence 8, Appli
37					Sequence 1, Appli
38					Sequence 14, Appli
39					Sequence 3, Appli
40					Sequence 3, Appli
41					Sequence 17, Appli
42					Sequence 8, Appli
43					Sequence 17371, A
44					Sequence 2954, Ap
45					Sequence 5001, Ap
46					Sequence 18848, A
47					Sequence 32367, A
48					Sequence 41438, A
49					Sequence 6, Appli
50					Sequence 6, Appli
51					Sequence 22, Appli
52					Sequence 13, Appli
53					Sequence 2856, Ap
54					Sequence 28, Appli
55					Sequence 28, Appli
56					Sequence 29, Appli
57					Sequence 29, Appli
58					Sequence 450, App
59					Sequence 22838, A
60					Sequence 4010, Ap
61					Sequence 23994, A
62					Sequence 40697, A
63					Sequence 55913, A
64					Sequence 2, Appli
65					Sequence 2, Appli
66					Sequence 2, Appli
67					Sequence 2, Appli
68					Sequence 2, Appli
69					Sequence 2, Appli
70					Sequence 2, Appli
71					Sequence 2, Appli
72					Sequence 2, Appli
73					Sequence 2, Appli
74					Sequence 2, Appli
75					Sequence 2, Appli
76					Sequence 2, Appli
77					Sequence 2, Appli
78					Sequence 2, Appli
79					Sequence 2, Appli
80					Sequence 2, Appli
81					Sequence 2, Appli
82					Sequence 2, Appli
83					Sequence 2, Appli
84					Sequence 2, Appli
85					Sequence 2, Appli
86					Sequence 2, Appli
87					Sequence 2, Appli
88					Sequence 2, Appli
89					Sequence 2, Appli
90					Sequence 2, Appli
91					Sequence 2, Appli
92					Sequence 2, Appli
93					Sequence 2, Appli
94					Sequence 2, Appli
95					Sequence 2, Appli
96					Sequence 2, Appli
97					Sequence 2, Appli
98					Sequence 2, Appli
99					Sequence 2, Appli
100					Sequence 2, Appli

Sequence 16, Appli  
Sequence 59, Appli  
Sequence 6, Appli  
Sequence 23, Appli  
Sequence 8, Appli  
Sequence 8, Appli  
Sequence 20, Appli  
Sequence 11, Appli  
Sequence 4, Appli  
Sequence 4, Appli  
Sequence 8048, Ap  
Sequence 7, Appli  
Sequence 7, Appli  
Sequence 21, Appli  
Sequence 12, Appli  
Sequence 18288, A  
Sequence 7, Appli  
Sequence 7, Appli  
Sequence 12703, A  
Sequence 20502, A  
Sequence 1, Appli  
Sequence 1, Appli  
Sequence 2, Appli  
Sequence 1, Appli  
Sequence 9, Appli  
Sequence 2, Appli  
Sequence 2, Appli  
Sequence 84, Appli  
Sequence 1, Appli  
Sequence 85, Appli  
Sequence 14, Appli  
Sequence 3, Appli  
Sequence 3, Appli  
Sequence 17, Appli  
Sequence 8, Appli  
Sequence 17371, A  
Sequence 2954, Ap  
Sequence 5001, Ap  
Sequence 18848, A  
Sequence 32367, A  
Sequence 41438, A  
Sequence 6, Appli  
Sequence 6, Appli  
Sequence 22, Appli  
Sequence 13, Appli  
Sequence 2856, Ap  
Sequence 28, Appli  
Sequence 28, Appli  
Sequence 29, Appli  
Sequence 29, Appli  
Sequence 450, App  
Sequence 22838, A  
Sequence 4010, Ap  
Sequence 23994, A  
Sequence 40697, A  
Sequence 55913, A  
Sequence 2, Appli  
Sequence 2, Appli  
Sequence 2, Appli  
Sequence 2, Appli  
Sequence 2, Appli  
Sequence 2, Appli  
Sequence 2, Appli  
Sequence 12535, A  
Sequence 17420, A  
Sequence 13960, A

101	37	54.4	388	2	US-08-382-505-2	Sequence 2, Appl1	174	34	50.0	38	1	US-08-451-472-40	Sequence 40, Appl
102	37	54.4	495	4	US-09-248-796A-16754	Sequence 16754, A	175	34	50.0	38	1	US-08-451-472-56	Sequence 56, Appl
103	37	54.4	702	3	US-09-232-200-102	Sequence 102, App	176	34	50.0	38	2	US-08-662-227-36	Sequence 36, Appl
104	37	54.4	702	3	US-09-232-197-102	Sequence 102, App	177	34	50.0	38	3	US-08-460-576-5	Sequence 5, Appl1
105	37	54.4	702	3	US-09-232-195-102	Sequence 102, App	178	34	50.0	38	3	US-09-017-947-36	Sequence 36, Appl
106	37	54.4	702	4	US-09-232-195-102	Sequence 102, App	179	34	50.0	38	4	US-09-925-442-36	Sequence 36, Appl
107	37	54.4	788	4	US-09-583-110-3939	Sequence 3939, Ap	180	34	50.0	73	1	US-08-451-472-58	Sequence 58, Appl
108	37	54.4	856	4	US-09-107-433-3534	Sequence 3534, Ap	181	34	50.0	73	1	US-08-451-472-64	Sequence 64, Appl
109	36	52.9	26	2	US-08-928-958-2	Sequence 2, Appl1	182	34	50.0	83	4	US-09-621-976-6783	Sequence 6783, Ap
110	36	52.9	26	2	US-09-072-429-2	Sequence 2, Appl1	183	34	50.0	108	1	US-08-451-472-23	Sequence 23, Appl
111	36	52.9	62	2	US-08-530-569B-21	Sequence 21, Appl	184	34	50.0	108	1	US-08-451-472-33	Sequence 33, Appl
112	36	52.9	122	3	US-08-858-207A-321	Sequence 321, App	185	34	50.0	113	4	US-09-902-540-11115	Sequence 11115, A
113	36	52.9	136	4	US-09-328-352-5712	Sequence 5712, Ap	186	34	50.0	117	4	US-09-248-796A-16571	Sequence 16571, A
114	36	52.9	141	4	US-09-489-039A-13440	Sequence 13440, A	187	34	50.0	131	1	US-08-154-916-2	Sequence 2, Appl1
115	36	52.9	208	4	US-09-583-110-4112	Sequence 4112, Ap	188	34	50.0	131	2	US-08-675-508-1	Sequence 1, Appl1
116	36	52.9	212	4	US-09-107-433-2936	Sequence 2936, Ap	189	34	50.0	131	2	US-09-139-424-2	Sequence 2, Appl1
117	36	52.9	267	4	US-10-000-489-80	Sequence 80, Appl	190	34	50.0	131	3	US-08-746-397-2	Sequence 2, Appl1
118	36	52.9	274	4	US-09-543-681A-8301	Sequence 8301, Ap	191	34	50.0	131	3	US-09-203-939-5	Sequence 5, Appl1
119	36	52.9	304	4	US-09-893-737-304	Sequence 304, App	192	34	50.0	131	3	US-09-251-835-5	Sequence 5, Appl1
120	36	52.9	328	4	US-09-902-540-10796	Sequence 10796, A	193	34	50.0	131	3	US-09-318-503-5	Sequence 5, Appl1
121	36	52.9	338	4	US-09-270-767-44254	Sequence 44254, A	194	34	50.0	131	3	US-09-038-261A-5	Sequence 5, Appl1
122	36	52.9	410	4	US-09-949-016-10345	Sequence 10345, A	195	34	50.0	131	4	US-09-564-329A-5	Sequence 5, Appl1
123	36	52.9	410	4	US-09-949-016-10345	Sequence 10345, A	196	34	50.0	131	4	US-09-963-620-5	Sequence 5, Appl1
124	36	52.9	716	4	US-09-270-767-41668	Sequence 41668, A	197	34	50.0	131	4	US-09-855-632-5	Sequence 5, Appl1
125	36	52.9	820	4	US-09-252-991A-29408	Sequence 29408, A	198	34	50.0	131	4	US-09-949-016-6211	Sequence 6211, Ap
126	36	52.9	886	4	US-09-902-540-11581	Sequence 11581, A	199	34	50.0	141	4	US-09-934-773-5	Sequence 5, Appl1
127	35	51.5	8	3	US-09-186-170-2	Sequence 2, Appl1	200	34	50.0	141	4	US-09-949-016-8741	Sequence 8741, Ap
128	35	51.5	8	4	US-09-562-868-2	Sequence 2, Appl1	201	34	50.0	173	3	US-09-093-855-207	Sequence 207, App
129	35	51.5	8	4	US-10-116-288A-2	Sequence 2, Appl1	202	34	50.0	173	3	US-09-205-426-207	Sequence 207, App
130	35	51.5	8	4	US-09-671-089-18	Sequence 18, Appl	203	34	50.0	199	4	US-09-270-767-41383	Sequence 41383, A
131	35	51.5	10	4	US-09-671-089-9	Sequence 9, Appl1	204	34	50.0	199	4	US-09-270-767-56599	Sequence 56599, A
132	35	51.5	40	4	US-09-902-540-14864	Sequence 14864, A	205	34	50.0	248	1	US-08-426-819A-37	Sequence 37, Appl
133	35	51.5	109	4	US-09-059-625-40	Sequence 40, Appl	206	34	50.0	258	4	US-09-252-991A-30990	Sequence 30990, A
134	35	51.5	122	4	US-09-252-991A-30865	Sequence 30865, A	207	34	50.0	270	3	US-09-323-872A-38	Sequence 38, Appl
135	35	51.5	136	2	US-08-675-508-5	Sequence 5, Appl1	208	34	50.0	289	4	US-09-489-039A-12390	Sequence 12390, A
136	35	51.5	138	3	US-08-746-397-11	Sequence 11, Appl	209	34	50.0	290	4	US-09-360-376-12	Sequence 12, Appl
137	35	51.5	140	4	US-09-270-767-47017	Sequence 47017, A	210	34	50.0	298	4	US-09-248-796A-19885	Sequence 19885, A
138	35	51.5	178	4	US-09-252-991A-20957	Sequence 20957, A	211	34	50.0	322	4	US-09-360-376-16	Sequence 16, Appl
139	35	51.5	181	4	US-09-252-991A-31827	Sequence 31827, A	212	34	50.0	329	4	US-09-252-991A-18420	Sequence 18420, A
140	35	51.5	192	4	US-09-252-991A-32913	Sequence 32913, A	213	34	50.0	332	4	US-09-270-767-42655	Sequence 42655, A
141	35	51.5	214	4	US-09-543-681A-6802	Sequence 6802, Ap	214	34	50.0	375	4	US-09-902-540-9768	Sequence 9768, Ap
142	35	51.5	355	4	US-09-252-991A-22494	Sequence 22494, A	215	34	50.0	380	4	US-09-902-540-14637	Sequence 14637, A
143	35	51.5	418	4	US-09-252-991A-17598	Sequence 17598, A	216	34	50.0	390	3	US-08-460-576-2	Sequence 2, Appl1
144	35	51.5	440	4	US-09-540-236-2067	Sequence 2067, Ap	217	34	50.0	395	4	US-09-489-039A-7739	Sequence 7739, Ap
145	35	51.5	460	4	US-09-949-016-7994	Sequence 7994, Ap	218	34	50.0	423	3	US-09-134-001C-5536	Sequence 5536, Ap
146	35	51.5	481	4	US-09-198-452A-1012	Sequence 1012, Ap	219	34	50.0	445	4	US-09-252-991A-18503	Sequence 18503, A
147	35	51.5	538	4	US-09-489-039A-8363	Sequence 8363, Ap	220	34	50.0	465	4	US-09-270-767-43452	Sequence 43452, A
148	35	51.5	539	4	US-09-134-000C-4363	Sequence 4363, Ap	221	34	50.0	468	4	US-09-602-787A-324	Sequence 324, App
149	35	51.5	578	4	US-09-949-016-11640	Sequence 11640, A	222	34	50.0	469	4	US-09-270-767-45555	Sequence 45555, A
150	35	51.5	617	4	US-09-902-540-14297	Sequence 14297, A	223	34	50.0	477	4	US-09-252-991A-27656	Sequence 27656, A
151	35	51.5	627	4	US-08-851-567B-28	Sequence 28, Appl	224	34	50.0	483	4	US-09-902-540-9804	Sequence 9804, Ap
152	35	51.5	644	4	US-09-252-991A-25084	Sequence 25084, A	225	34	50.0	530	3	US-09-103-039A-4	Sequence 4, Appl1
153	35	51.5	659	4	US-09-543-681A-4748	Sequence 4748, Ap	226	34	50.0	539	3	US-08-941-445A-21	Sequence 21, Appl
154	35	51.5	802	4	US-09-252-991A-23824	Sequence 23824, A	227	34	50.0	580	4	US-09-949-016-10022	Sequence 10022, A
155	35	51.5	879	4	US-09-107-532A-4679	Sequence 4679, Ap	228	34	50.0	583	3	US-08-941-445A-13	Sequence 13, Appl
156	35	51.5	903	4	US-09-270-767-46470	Sequence 46470, A	229	34	50.0	584	1	US-08-426-819A-36	Sequence 36, Appl
157	35	51.5	1044	4	US-09-252-991A-18853	Sequence 18853, A	230	34	50.0	592	2	US-08-946-412-2	Sequence 2, Appl1
158	35	51.5	1189	4	US-08-851-567B-26	Sequence 26, Appl	231	34	50.0	604	3	US-09-391-104-30	Sequence 30, Appl
159	35	51.5	1319	2	US-08-290-731C-2	Sequence 2, Appl1	232	34	50.0	607	3	US-09-000-041A-2	Sequence 2, Appl1
160	35	51.5	1333	3	US-09-356-952-2	Sequence 2, Appl1	233	34	50.0	607	3	US-09-211-704A-10	Sequence 10, Appl
161	35	51.5	1336	4	US-09-976-594-312	Sequence 312, App	234	34	50.0	607	4	US-09-734-002-2	Sequence 2, Appl1
162	35	51.5	1336	2	US-08-290-731C-6	Sequence 6, Appl1	235	34	50.0	617	1	US-09-252-991A-21113	Sequence 21113, A
163	34	50.0	21	2	US-08-472-172-10	Sequence 10, Appl	236	34	50.0	622	1	US-08-426-819A-35	Sequence 35, Appl
164	34	50.0	27	2	US-08-928-958-3	Sequence 3, Appl1	237	34	50.0	622	4	US-09-513-597A-21	Sequence 21, Appl
165	34	50.0	27	2	US-09-072-429-3	Sequence 3, Appl1	238	34	50.0	622	4	US-09-949-016-6660	Sequence 6660, Ap
166	34	50.0	29	2	US-08-928-958-24	Sequence 24, Appl	239	34	50.0	628	3	US-08-776-271-2	Sequence 2, Appl1
167	34	50.0	29	2	US-09-072-429-24	Sequence 24, Appl	240	34	50.0	628	3	US-09-215-035-2	Sequence 2, Appl1
168	34	50.0	30	2	US-08-928-958-1	Sequence 1, Appl1	241	34	50.0	628	4	US-09-513-597A-20	Sequence 20, Appl
169	34	50.0	31	2	US-09-072-429-1	Sequence 1, Appl1	242	34	50.0	628	4	US-09-949-016-6849	Sequence 6849, Ap
170	34	50.0	31	2	US-08-572-951-25	Sequence 25, Appl	243	34	50.0	649	3	US-09-192-909-2	Sequence 2, Appl1
171	34	50.0	33	2	US-08-572-951-5	Sequence 5, Appl1	244	34	50.0	649	4	US-09-931-297-2	Sequence 2, Appl1
172	34	50.0	36	4	US-09-270-767-56911	Sequence 56911, A	245	34	50.0	659	4	US-09-949-016-7616	Sequence 7616, Ap
173	34	50.0	38	1	US-08-451-472-11	Sequence 11, Appl	246	34	50.0	659	4	US-09-949-016-7617	Sequence 7617, Ap

247	34	50.0	660	3	US-09-058-489-18	Sequence 18, Appl	320	48.5	502	4	US-09-724-797-90	Sequence 90, Appl
248	34	50.0	660	3	US-09-058-489-91	Sequence 91, Appl	321	48.5	524	4	US-09-903-540-16272	Sequence 16272, A
249	34	50.0	660	3	US-09-976-594-787	Sequence 787, App	322	48.5	531	4	US-09-253-991A-22509	Sequence 22509, A
250	34	50.0	662	3	US-09-058-489-15	Sequence 15, Appl	323	48.5	541	4	US-09-252-991A-27169	Sequence 27169, A
251	34	50.0	662	3	US-09-058-489-16	Sequence 16, Appl	324	48.5	585	4	US-09-252-991A-17061	Sequence 17061, A
252	34	50.0	662	4	US-09-856-930-6	Sequence 6, Appl	325	48.5	590	4	US-09-538-092-967	Sequence 967, App
253	34	50.0	666	4	US-09-489-039A-12388	Sequence 12388, A	326	48.5	629	4	US-09-603-787A-624	Sequence 624, App
254	34	50.0	667	4	US-09-747-259-16	Sequence 16, Appl	327	48.5	641	3	US-08-836-567-10	Sequence 10, Appl
255	34	50.0	667	4	US-09-816-744-16	Sequence 16, Appl	328	48.5	641	3	US-09-606-304-10	Sequence 10, Appl
256	34	50.0	671	3	US-09-196-390-2	Sequence 2, Appl	329	48.5	678	4	US-09-196-270-1	Sequence 1, Appl
257	34	50.0	671	4	US-09-952-677-2	Sequence 2, Appl	330	48.5	692	4	US-09-252-991A-19668	Sequence 19668, A
258	34	50.0	729	4	US-09-949-016-10790	Sequence 10790, A	331	48.5	724	4	US-09-949-016-10086	Sequence 10086, A
259	34	50.0	729	4	US-09-949-016-10791	Sequence 10791, A	332	48.5	777	4	US-09-903-540-15059	Sequence 15059, A
260	34	50.0	1093	4	US-09-252-991A-21827	Sequence 21827, A	333	48.5	815	4	US-09-196-270-7	Sequence 7, Appl
261	34	50.0	1296	4	US-09-462-136-9	Sequence 9, Appl	334	48.5	969	4	US-09-253-991A-26985	Sequence 26985, A
262	34	50.0	1607	4	US-09-902-540-16765	Sequence 16765, A	335	48.5	1025	4	US-09-248-796A-20573	Sequence 20573, A
263	33	48.5	36	4	US-09-914-259-87	Sequence 87, Appl	336	48.5	1128	4	US-09-252-991A-19672	Sequence 19672, A
264	33	48.5	66	4	US-09-489-039A-10473	Sequence 10473, A	337	48.5	1162	4	US-09-253-991A-32764	Sequence 32764, A
265	33	48.5	83	4	US-09-489-039A-7832	Sequence 7832, Ap	338	48.5	1195	4	US-09-538-092-517	Sequence 517, App
266	33	48.5	97	4	US-09-252-991A-24855	Sequence 24855, A	339	48.5	1587	4	US-09-845-583A-10	Sequence 10, Appl
267	33	48.5	101	4	US-09-252-991A-25246	Sequence 25246, A	340	48.5	1587	4	US-09-561-709B-3	Sequence 3, Appl
268	33	48.5	156	4	US-09-270-767-38035	Sequence 38035, A	341	48.5	1627	1	US-07-665-792E-9	Sequence 9, Appl
269	33	48.5	156	4	US-09-270-767-53252	Sequence 53252, A	342	48.5	1838	4	US-09-270-767-45188	Sequence 45188, A
270	33	48.5	219	4	US-09-050-739-14	Sequence 14, Appl	343	32.5	166	4	US-09-903-540-13098	Sequence 13098, A
271	33	48.5	223	4	US-09-252-991A-20177	Sequence 20177, A	344	32.5	414	4	US-09-252-991A-17176	Sequence 17176, A
272	33	48.5	232	4	US-09-107-532A-6091	Sequence 6091, Ap	345	47.8	427	4	US-09-543-681A-6225	Sequence 6225, Ap
273	33	48.5	238	4	US-09-252-991A-20037	Sequence 20037, A	346	47.1	14	2	US-08-503-226B-41	Sequence 41, Appl
274	33	48.5	248	4	US-09-543-681A-5849	Sequence 5849, Ap	347	47.1	14	3	US-08-721-458B-41	Sequence 41, Appl
275	33	48.5	259	4	US-09-328-352-6031	Sequence 6031, Ap	348	47.1	23	3	US-07-927-391-7	Sequence 7, Appl
276	33	48.5	260	4	US-09-902-540-14188	Sequence 14188, A	349	47.1	29	3	US-07-927-391-8	Sequence 8, Appl
277	33	48.5	267	4	US-09-248-796A-14590	Sequence 14590, A	350	47.1	31	3	US-08-996-139-9	Sequence 9, Appl
278	33	48.5	267	4	US-09-902-540-14651	Sequence 14651, A	351	47.1	31	3	US-08-995-659-9	Sequence 9, Appl
279	33	48.5	271	4	US-09-902-540-15639	Sequence 15639, A	352	47.1	31	3	US-09-215-649A-9	Sequence 9, Appl
280	33	48.5	276	1	US-08-467-155A-1	Sequence 1, Appl	353	47.1	31	4	US-08-577-780-9	Sequence 9, Appl
281	33	48.5	276	2	US-08-628-198-1	Sequence 1, Appl	354	47.1	31	4	US-09-577-800-9	Sequence 9, Appl
282	33	48.5	276	3	US-09-201-038-1	Sequence 1, Appl	355	47.1	31	4	US-09-466-496-9	Sequence 9, Appl
283	33	48.5	276	5	PCT-US96-07343-1	Sequence 1, Appl	356	47.1	31	4	US-09-871-856-9	Sequence 9, Appl
284	33	48.5	290	4	US-09-543-681A-4737	Sequence 4737, Ap	357	47.1	31	4	US-09-871-291-9	Sequence 9, Appl
285	33	48.5	291	1	US-08-467-155A-11	Sequence 11, Appl	358	47.1	31	4	US-09-877-650-9	Sequence 9, Appl
286	33	48.5	291	2	US-08-628-198-11	Sequence 11, Appl	359	47.1	31	4	US-08-865-383-9	Sequence 9, Appl
287	33	48.5	291	3	US-09-201-038-11	Sequence 11, Appl	360	47.1	31	4	US-09-688-459-9	Sequence 9, Appl
288	33	48.5	291	5	PCT-US96-07343-11	Sequence 11, Appl	361	47.1	32	4	US-08-670-354-9	Sequence 9, Appl
289	33	48.5	292	4	US-09-543-681A-6495	Sequence 6495, Ap	362	47.1	32	3	US-09-320-424-9	Sequence 9, Appl
290	33	48.5	311	4	US-09-602-787A-28	Sequence 28, Appl	363	47.1	32	3	PCT-US96-10895-9	Sequence 9, Appl
291	33	48.5	314	4	US-09-360-376-14	Sequence 14, Appl	364	47.1	32	5	US-07-927-391-9	Sequence 9, Appl
292	33	48.5	342	1	US-08-118-270-9	Sequence 9, Appl	365	47.1	32	3	US-09-227-357-654	Sequence 654, App
293	33	48.5	342	5	PCT-US93-08528-9	Sequence 9, Appl	366	47.1	66	3	US-09-270-767-37426	Sequence 37426, A
294	33	48.5	344	4	US-09-252-991A-30554	Sequence 30554, A	367	47.1	72	4	US-08-858-207A-476	Sequence 476, App
295	33	48.5	354	1	US-08-313-553-9	Sequence 9, Appl	368	47.1	72	4	US-08-204-740-13	Sequence 13, Appl
296	33	48.5	354	3	US-08-767-993-9	Sequence 11, Appl	369	47.1	81	3	US-08-081-167A-13	Sequence 13, Appl
297	33	48.5	355	1	US-08-118-270-11	Sequence 11, Appl	370	47.1	87	3	US-09-081-395-13	Sequence 13, Appl
298	33	48.5	355	5	PCT-US93-08528-11	Sequence 11, Appl	371	47.1	87	3	US-09-416-833-13	Sequence 13, Appl
299	33	48.5	362	4	US-09-540-236-2243	Sequence 2243, Ap	372	47.1	87	3	PCT-US95-02521-13	Sequence 13, Appl
300	33	48.5	371	4	US-09-540-236-3776	Sequence 3776, Ap	373	47.1	87	5	US-09-543-681A-8178	Sequence 8178, Ap
301	33	48.5	373	4	US-09-602-787A-528	Sequence 528, App	374	47.1	95	4	US-08-480-449-18	Sequence 18, Appl
302	33	48.5	382	4	US-09-711-164-399	Sequence 399, App	375	47.1	99	1	US-08-660-542-18	Sequence 18, Appl
303	33	48.5	387	4	US-09-180-109A-2	Sequence 2, Appl	376	47.1	99	2	US-08-613-822-18	Sequence 18, Appl
304	33	48.5	387	4	US-09-180-109A-6	Sequence 6, Appl	377	47.1	99	3	US-08-479-729B-18	Sequence 18, Appl
305	33	48.5	387	4	US-09-180-109A-8	Sequence 8, Appl	378	47.1	99	3	US-08-939-107-18	Sequence 22, Appl
306	33	48.5	397	4	US-09-602-787A-628	Sequence 628, App	379	47.1	99	4	US-09-717-209-18	Sequence 18, Appl
307	33	48.5	426	4	US-09-489-039A-11995	Sequence 11995, A	380	47.1	99	4	US-09-545-894-7	Sequence 7, Appl
308	33	48.5	428	4	US-09-252-991A-19535	Sequence 19535, A	381	47.1	99	4	US-09-834-795A-27	Sequence 27, Appl
309	33	48.5	439	4	US-09-252-991A-21998	Sequence 21998, A	382	47.1	99	4	US-09-067-447B-18	Sequence 18, Appl
310	33	48.5	444	4	US-09-248-796A-18129	Sequence 18129, A	383	47.1	99	4	US-08-920-821-2	Sequence 2, Appl
311	33	48.5	450	4	US-09-248-796A-15183	Sequence 15183, A	384	47.1	104	3	US-09-107-433-2648	Sequence 2648, Ap
312	33	48.5	459	3	US-08-836-567-4	Sequence 4, Appl	385	47.1	108	1	US-08-204-740-11	Sequence 11, Appl
313	33	48.5	459	4	US-09-606-304-4	Sequence 4, Appl	386	47.1	108	3	US-09-081-167A-11	Sequence 11, Appl
314	33	48.5	461	1	US-08-196-989B-13	Sequence 13, Appl	387	47.1	104	3		
315	33	48.5	461	2	US-08-760-936-13	Sequence 13, Appl	388	47.1	104	3		
316	33	48.5	461	4	US-09-225-024-13	Sequence 13, Appl	389	47.1	104	3		
317	33	48.5	473	4	US-09-949-016-7944	Sequence 7944, Ap	390	47.1	107	4		
318	33	48.5	479	1	US-08-313-553-7	Sequence 7, Appl	391	47.1	108	1		
319	33	48.5	479	3	US-08-767-993-7	Sequence 7, Appl	392	47.1	108	3		

393	32	47.1	108	3	US-09-081-395-11	Sequence 11, Appl	466	32	47.1	411	3	US-08-981-189B-12	Sequence 12, Appl
394	32	47.1	108	3	US-09-416-833-11	Sequence 11, Appl	467	32	47.1	411	4	US-09-799-978-10	Sequence 10, Appl
395	32	47.1	108	5	PCI-US95-02541-11	Sequence 11, Appl	468	32	47.1	411	4	US-09-799-978-18	Sequence 18, Appl
396	32	47.1	109	2	US-08-421-144A-7	Sequence 7, Appl	469	32	47.1	411	4	US-09-881-401-4	Sequence 4, Appl
397	32	47.1	109	3	US-07-927-391-16	Sequence 16, Appl	470	32	47.1	411	4	US-09-881-401-8	Sequence 8, Appl
398	32	47.1	109	3	US-08-679-493A-153	Sequence 153, Appl	471	32	47.1	411	4	US-09-631-603-12	Sequence 12, Appl
399	32	47.1	110	4	US-09-366-887A-13	Sequence 13, Appl	472	32	47.1	413	4	US-09-799-978-32	Sequence 32, Appl
400	32	47.1	110	4	US-09-517-204-13	Sequence 13, Appl	473	32	47.1	413	4	US-08-110-286A-2	Sequence 2, Appl
401	32	47.1	115	4	US-09-774-639-154	Sequence 154, Appl	474	32	47.1	415	1	US-08-110-286A-6	Sequence 6, Appl
402	32	47.1	122	4	US-09-949-016-9711	Sequence 9711, Ap	475	32	47.1	415	1	US-08-981-189B-10	Sequence 10, Appl
403	32	47.1	126	4	US-09-489-039A-11199	Sequence 11199, A	476	32	47.1	415	3	US-08-482-746-2	Sequence 2, Appl
404	32	47.1	129	4	US-09-328-352-4959	Sequence 4959, Ap	477	32	47.1	415	3	US-08-482-746-6	Sequence 6, Appl
405	32	47.1	129	4	US-09-302-540-13469	Sequence 13469, A	478	32	47.1	415	3	US-08-482-746-13	Sequence 13, Appl
406	32	47.1	135	4	US-09-894-998A-25	Sequence 25, Appl	479	32	47.1	415	4	US-09-580-734-2	Sequence 2, Appl
407	32	47.1	135	4	US-09-270-767-37619	Sequence 37619, A	480	32	47.1	415	4	US-09-580-734-6	Sequence 6, Appl
408	32	47.1	135	4	US-09-270-767-52836	Sequence 52836, A	481	32	47.1	415	4	US-09-580-734-13	Sequence 13, Appl
409	32	47.1	135	4	US-10-237-551-25	Sequence 25, Appl	482	32	47.1	415	4	US-08-374-009-2	Sequence 2, Appl
410	32	47.1	136	4	US-09-252-991A-20815	Sequence 20815, A	483	32	47.1	415	4	US-08-374-009-6	Sequence 6, Appl
411	32	47.1	137	4	US-10-081-817A-30	Sequence 30, Appl	484	32	47.1	415	4	US-08-374-009-13	Sequence 13, Appl
412	32	47.1	146	4	US-09-583-110-4509	Sequence 4509, Ap	485	32	47.1	415	4	US-09-191-724-2	Sequence 2, Appl
413	32	47.1	151	4	US-09-502-540-14278	Sequence 14278, A	486	32	47.1	415	4	US-09-191-724-6	Sequence 6, Appl
414	32	47.1	159	4	US-09-252-991A-29207	Sequence 29207, A	487	32	47.1	415	4	US-09-191-724-13	Sequence 13, Appl
415	32	47.1	169	4	US-09-540-236-3364	Sequence 3364, Ap	488	32	47.1	415	4	US-09-799-978-2	Sequence 2, Appl
416	32	47.1	170	4	US-09-270-767-47342	Sequence 47342, A	489	32	47.1	415	4	US-09-799-978-4	Sequence 4, Appl
417	32	47.1	180	3	US-09-449-218D-44	Sequence 44, Appl	490	32	47.1	415	4	US-09-799-978-16	Sequence 16, Appl
418	32	47.1	180	4	US-09-668-529A-44	Sequence 44, Appl	491	32	47.1	415	4	US-09-799-978-22	Sequence 22, Appl
419	32	47.1	180	4	US-09-668-037A-44	Sequence 44, Appl	492	32	47.1	415	4	US-09-799-978-28	Sequence 28, Appl
420	32	47.1	181	4	US-09-040-229B-8	Sequence 8, Appl	493	32	47.1	415	4	US-09-799-978-30	Sequence 30, Appl
421	32	47.1	187	6	5217891-4	Patent No. 5217891	494	32	47.1	415	4	US-09-799-978-40	Sequence 40, Appl
422	32	47.1	187	6	5217891-4	Patent No. 5217891	495	32	47.1	422	4	US-09-252-991A-25540	Sequence 25540, A
423	32	47.1	196	4	US-09-543-681A-7659	Sequence 7659, Ap	496	32	47.1	425	4	US-09-252-991A-27848	Sequence 27848, A
424	32	47.1	207	4	US-08-811-519-16	Sequence 16, Appl	497	32	47.1	426	4	US-09-602-787A-424	Sequence 424, Appl
425	32	47.1	207	4	US-08-811-519-17	Sequence 17, Appl	498	32	47.1	428	4	US-09-799-978-36	Sequence 36, Appl
426	32	47.1	207	4	US-08-811-519-18	Sequence 18, Appl	499	32	47.1	430	3	US-09-799-978-26	Sequence 26, Appl
427	32	47.1	208	4	US-09-569-611C-27	Sequence 27, Appl	500	32	47.1	430	3	US-09-799-978-26	Sequence 26, Appl
428	32	47.1	210	4	US-09-569-611C-28	Sequence 28, Appl	501	32	47.1	431	1	US-08-381-189B-11	Sequence 11, Appl
429	32	47.1	222	4	US-09-252-991A-30827	Sequence 30827, A	502	32	47.1	431	3	US-08-981-189B-13	Sequence 13, Appl
430	32	47.1	239	3	US-09-372-422A-42	Sequence 42, Appl	503	32	47.1	431	3	US-08-482-746-10	Sequence 10, Appl
431	32	47.1	256	3	US-09-320-424-13	Sequence 13, Appl	504	32	47.1	431	3	US-09-580-734-10	Sequence 10, Appl
432	32	47.1	256	4	US-09-825-563-13	Sequence 13, Appl	505	32	47.1	431	4	US-09-580-734-10	Sequence 10, Appl
433	32	47.1	271	4	US-09-902-540-9954	Sequence 9954, Ap	506	32	47.1	431	4	US-08-374-009-10	Sequence 10, Appl
434	32	47.1	276	4	US-09-489-039A-9409	Sequence 9409, Ap	507	32	47.1	431	4	US-09-191-724-10	Sequence 10, Appl
435	32	47.1	283	4	US-09-145-828A-22	Sequence 22, Appl	508	32	47.1	431	4	US-09-799-978-24	Sequence 24, Appl
436	32	47.1	283	4	US-09-903-456-29	Sequence 29, Appl	509	32	47.1	431	4	US-09-881-401-2	Sequence 2, Appl
437	32	47.1	291	4	US-09-252-991A-28294	Sequence 28294, A	510	32	47.1	438	4	US-09-799-978-12	Sequence 12, Appl
438	32	47.1	307	4	US-09-489-039A-8974	Sequence 8974, Ap	511	32	47.1	440	4	US-09-631-603-22	Sequence 22, Appl
439	32	47.1	310	2	US-08-943-600A-3	Sequence 3, Appl	512	32	47.1	440	4	US-09-826-509-567	Sequence 567, Appl
440	32	47.1	310	4	US-09-354-942-3	Sequence 3, Appl	513	32	47.1	444	3	US-08-482-746-15	Sequence 15, Appl
441	32	47.1	313	4	US-09-248-796A-17621	Sequence 17621, A	514	32	47.1	444	4	US-09-580-734-15	Sequence 15, Appl
442	32	47.1	317	4	US-09-540-236-3790	Sequence 3790, Ap	515	32	47.1	444	4	US-08-374-009-15	Sequence 15, Appl
443	32	47.1	318	3	US-08-680-506-3	Sequence 3, Appl	516	32	47.1	444	4	US-09-191-724-15	Sequence 15, Appl
444	32	47.1	332	4	US-09-248-796A-18143	Sequence 18143, A	517	32	47.1	444	4	US-09-631-603-11	Sequence 11, Appl
445	32	47.1	337	4	US-09-543-681A-6463	Sequence 6463, Ap	518	32	47.1	445	4	US-09-799-978-34	Sequence 34, Appl
446	32	47.1	353	4	US-09-902-540-13000	Sequence 13000, A	519	32	47.1	445	4	US-09-902-540-13038	Sequence 13038, A
447	32	47.1	355	4	US-09-498-520A-28	Sequence 28, Appl	520	32	47.1	447	4	US-09-489-039A-11893	Sequence 11893, A
448	32	47.1	365	4	US-09-252-991A-21955	Sequence 21955, A	521	32	47.1	449	3	US-08-680-506-7	Sequence 7, Appl
449	32	47.1	365	4	US-09-248-796A-21739	Sequence 21739, A	522	32	47.1	451	4	US-09-949-016-7978	Sequence 7978, Ap
450	32	47.1	369	4	US-09-603-208A-164	Sequence 164, App	523	32	47.1	451	4	US-09-949-016-7979	Sequence 7979, Ap
451	32	47.1	375	4	US-09-799-978-6	Sequence 6, Appl	524	32	47.1	451	4	US-09-949-016-7980	Sequence 7980, Ap
452	32	47.1	378	4	US-09-949-016-8375	Sequence 8375, Ap	525	32	47.1	451	4	US-09-949-016-7981	Sequence 7981, Ap
453	32	47.1	384	4	US-09-603-208A-162	Sequence 162, App	526	32	47.1	451	4	US-09-949-016-7982	Sequence 7982, Ap
454	32	47.1	388	4	US-09-583-110-3393	Sequence 3393, Ap	527	32	47.1	458	6	5217891-15	Patent No. 5217891
455	32	47.1	390	4	US-09-107-532A-6569	Sequence 6569, Ap	528	32	47.1	458	6	5217891-15	Patent No. 5217891
456	32	47.1	392	4	US-09-489-039A-14298	Sequence 14298, A	529	32	47.1	460	2	US-08-677-049-10	Sequence 10, Appl
457	32	47.1	393	4	US-09-489-039A-13122	Sequence 13122, A	530	32	47.1	461	4	US-09-489-039A-12176	Sequence 12176, A
458	32	47.1	395	4	US-09-248-796A-23772	Sequence 23772, A	531	32	47.1	465	4	US-09-252-991A-17635	Sequence 17635, A
459	32	47.1	396	4	US-09-107-433-4213	Sequence 4213, Ap	532	32	47.1	466	4	US-09-902-540-14180	Sequence 14180, A
460	32	47.1	397	4	US-09-799-978-14	Sequence 14, Appl	533	32	47.1	467	4	US-09-543-681A-7427	Sequence 7427, Ap
461	32	47.1	401	4	US-09-799-978-8	Sequence 8, Appl	534	32	47.1	471	4	US-09-252-991A-22393	Sequence 22393, A
462	32	47.1	405	4	US-09-799-978-38	Sequence 38, Appl	535	32	47.1	474	4	US-09-949-016-7983	Sequence 7983, Ap
463	32	47.1	410	4	US-09-603-208A-154	Sequence 154, App	536	32	47.1	474	4	US-09-949-016-7984	Sequence 7984, Ap
464	32	47.1	411	1	US-08-381-433A-4	Sequence 4, Appl	537	32	47.1	474	4	US-09-949-016-7985	Sequence 7985, Ap
465	32	47.1	411	1	US-08-381-433A-8	Sequence 8, Appl	538	32	47.1	474	4	US-09-949-016-7986	Sequence 7986, Ap



539	32	47.1	474	4	US-09-949-016-7987	Sequence 7987, Ap	612	31.5	46.3	480	3	US-08-923-454A-18	Sequence 18, Appl
540	32	47.1	476	4	US-09-328-352-4782	Sequence 4782, Ap	613	31.5	46.3	480	4	US-09-949-016-6270	Sequence 6270, Ap
541	32	47.1	501	4	US-09-902-540-11195	Sequence 11195, A	614	31.5	46.3	718	4	US-09-252-991A-25683	Sequence 25683, A
542	32	47.1	504	4	US-09-543-681A-6790	Sequence 6790, Ap	615	31	45.6	15	2	US-08-928-958-8	Sequence 8, Appl
543	32	47.1	509	4	US-09-602-787A-282	Sequence 282, Ap	616	31	45.6	15	2	US-08-928-958-8	Sequence 8, Appl
544	32	47.1	514	3	US-09-066-047-3	Sequence 3, Appl	617	31	45.6	16	1	US-08-642-493-1	Sequence 1, Appl
545	32	47.1	524	4	US-09-252-991A-18092	Sequence 18092, A	618	31	45.6	16	1	US-08-258-852-5	Sequence 5, Appl
546	32	47.1	531	3	US-09-380-420C-2	Sequence 2, Appl	619	31	45.6	16	2	US-08-928-958-4	Sequence 4, Appl
547	32	47.1	531	4	US-09-949-642A-2	Sequence 2, Appl	620	31	45.6	16	2	US-09-072-429-4	Sequence 4, Appl
548	32	47.1	552	4	US-09-949-016-11728	Sequence 11728, A	621	31	45.6	16	3	US-08-964-302A-3	Sequence 3, Appl
549	32	47.1	558	4	US-09-902-540-14971	Sequence 14971, A	622	31	45.6	16	3	US-08-170-754B-5	Sequence 5, Appl
550	32	47.1	576	4	US-09-543-681A-7562	Sequence 7562, Ap	623	31	45.6	16	3	US-09-441-416A-3	Sequence 3, Appl
551	32	47.1	595	4	US-10-028-272-5	Sequence 5, Appl	624	31	45.6	16	3	US-09-411-706-1	Sequence 1, Appl
552	32	47.1	596	4	US-09-949-016-6911	Sequence 6911, Ap	625	31	45.6	16	3	US-09-230-548-19	Sequence 19, Appl
553	32	47.1	596	4	US-09-949-016-9814	Sequence 9814, Ap	626	31	45.6	16	4	US-09-450-071A-5	Sequence 5, Appl
554	32	47.1	599	4	US-09-949-016-6008	Sequence 6008, Ap	627	31	45.6	16	4	US-09-935-032-1	Sequence 1, Appl
555	32	47.1	603	4	US-09-252-991A-31905	Sequence 31905, A	628	31	45.6	16	4	US-08-720-039-1	Sequence 1, Appl
556	32	47.1	604	4	US-09-640-419C-23	Sequence 23, Appl	629	31	45.6	16	4	US-10-083-889-7	Sequence 7, Appl
557	32	47.1	605	4	US-09-540-236-2056	Sequence 2056, Ap	630	31	45.6	16	4	US-10-116-288A-23	Sequence 23, Appl
558	32	47.1	631	4	US-09-252-991A-27902	Sequence 27902, A	631	31	45.6	16	4	US-10-144-549-18	Sequence 18, Appl
559	32	47.1	688	4	US-09-902-540-12710	Sequence 12710, A	632	31	45.6	16	5	PCT-US95-07539-5	Sequence 5, Appl
560	32	47.1	719	4	US-09-902-540-11526	Sequence 11526, A	633	31	45.6	20	1	US-08-399-696-100	Sequence 100, Ap
561	32	47.1	720	4	US-09-252-991A-19581	Sequence 19581, A	634	31	45.6	22	4	US-08-692-170C-24	Sequence 24, Appl
562	32	47.1	722	4	US-10-028-272-3	Sequence 3, Appl	635	31	45.6	22	4	US-10-405-231A-24	Sequence 24, Appl
563	32	47.1	723	4	US-10-028-272-2	Sequence 2, Appl	636	31	45.6	22	4	US-09-984-365-24	Sequence 24, Appl
564	32	47.1	726	4	US-09-489-039A-7465	Sequence 7465, Ap	637	31	45.6	22	4	US-08-399-696-97	Sequence 97, Appl
565	32	47.1	734	4	US-09-894-988A-26	Sequence 26, Appl	638	31	45.6	24	3	US-08-230-548-21	Sequence 21, Appl
566	32	47.1	734	4	US-10-237-551-25	Sequence 25, Appl	639	31	45.6	24	3	US-08-230-548-21	Sequence 21, Appl
567	32	47.1	743	4	US-10-028-272-4	Sequence 4, Appl	640	31	45.6	25	3	US-09-230-548-20	Sequence 20, Appl
568	32	47.1	760	4	US-09-252-991A-18711	Sequence 18711, A	641	31	45.6	26	1	US-08-258-852-1	Sequence 1, Appl
569	32	47.1	761	4	US-09-949-016-8473	Sequence 8473, Ap	642	31	45.6	26	1	US-08-258-852-3	Sequence 3, Appl
570	32	47.1	761	4	US-09-949-016-8474	Sequence 8474, Ap	643	31	45.6	26	1	US-08-258-852-8	Sequence 8, Appl
571	32	47.1	761	4	US-09-949-016-8475	Sequence 8475, Ap	644	31	45.6	26	1	US-08-258-852-9	Sequence 9, Appl
572	32	47.1	761	4	US-09-949-016-8476	Sequence 8476, Ap	645	31	45.6	26	1	US-08-258-852-9	Sequence 9, Appl
573	32	47.1	779	4	US-09-949-016-10692	Sequence 10692, A	646	31	45.6	26	2	US-08-928-958-5	Sequence 5, Appl
574	32	47.1	779	4	US-09-949-016-10693	Sequence 10693, A	647	31	45.6	26	2	US-08-928-958-5	Sequence 5, Appl
575	32	47.1	812	4	US-09-919-039-58	Sequence 58, Appl	648	31	45.6	26	3	US-09-170-754B-1	Sequence 1, Appl
576	32	47.1	836	4	US-09-949-016-11057	Sequence 11057, A	649	31	45.6	26	3	US-09-170-754B-8	Sequence 8, Appl
577	32	47.1	836	4	US-09-949-016-11058	Sequence 11058, A	650	31	45.6	26	3	US-09-170-754B-9	Sequence 9, Appl
578	32	47.1	836	4	US-09-949-016-11059	Sequence 11059, A	651	31	45.6	26	3	US-09-170-754B-9	Sequence 9, Appl
579	32	47.1	836	4	US-09-949-016-11060	Sequence 11060, A	652	31	45.6	26	4	US-09-450-071A-1	Sequence 1, Appl
580	32	47.1	1045	4	US-09-949-016-8469	Sequence 8469, Ap	653	31	45.6	26	4	US-09-450-071A-8	Sequence 8, Appl
581	32	47.1	1045	4	US-09-949-016-8470	Sequence 8470, Ap	654	31	45.6	26	4	US-09-450-071A-8	Sequence 8, Appl
582	32	47.1	1045	4	US-09-949-016-8471	Sequence 8471, Ap	655	31	45.6	26	4	US-09-450-071A-13	Sequence 13, Appl
583	32	47.1	1045	4	US-09-949-016-8472	Sequence 8472, Ap	656	31	45.6	26	4	US-09-720-039-3	Sequence 3, Appl
584	32	47.1	1160	4	US-09-711-164-401	Sequence 401, Ap	657	31	45.6	26	4	US-09-927-465B-1	Sequence 1, Appl
585	32	47.1	1203	3	US-09-207-857-2	Sequence 2, Appl	658	31	45.6	26	5	PCT-US95-07539-1	Sequence 1, Appl
586	32	47.1	1203	3	US-09-293-505-2	Sequence 2, Appl	659	31	45.6	26	5	PCT-US95-07539-3	Sequence 3, Appl
587	32	47.1	1203	3	US-09-909-280A-2	Sequence 2, Appl	660	31	45.6	26	5	PCT-US95-07539-8	Sequence 8, Appl
588	32	47.1	1203	3	US-09-060-939A-2	Sequence 2, Appl	661	31	45.6	26	5	PCT-US95-07539-9	Sequence 9, Appl
589	32	47.1	1216	4	US-09-248-796A-19646	Sequence 19646, A	662	31	45.6	28	1	US-08-258-852-4	Sequence 4, Appl
590	32	47.1	1296	4	US-08-857-636-60	Sequence 60, Appl	663	31	45.6	28	3	US-08-170-754B-4	Sequence 4, Appl
591	32	47.1	1434	2	US-08-540-406-10	Sequence 10, Appl	664	31	45.6	28	4	US-09-450-071A-4	Sequence 4, Appl
592	32	47.1	1434	3	US-08-656-055-10	Sequence 10, Appl	665	31	45.6	28	5	PCT-US95-07539-4	Sequence 4, Appl
593	32	47.1	1434	3	US-08-954-668-10	Sequence 10, Appl	666	31	45.6	30	3	US-09-051-934-53	Sequence 53, Appl
594	32	47.1	1434	3	US-08-918-658-10	Sequence 10, Appl	667	31	45.6	30	3	US-09-051-934-53	Sequence 53, Appl
595	32	47.1	1434	4	US-09-724-631-10	Sequence 10, Appl	668	31	45.6	30	4	US-03-471-276-1434	Sequence 1434, Ap
596	32	47.1	1434	4	US-08-954-701A-10	Sequence 10, Appl	669	31	45.6	36	3	US-08-082-279B-843	Sequence 843, Ap
597	32	47.1	1434	5	PCT-US95-13233-10	Sequence 10, Appl	670	31	45.6	36	3	US-08-082-279B-844	Sequence 844, Ap
598	32	47.1	1447	3	US-08-540-406-19	Sequence 19, Appl	671	31	45.6	36	3	US-09-315-304B-843	Sequence 843, Ap
599	32	47.1	1447	3	US-08-656-055-19	Sequence 19, Appl	672	31	45.6	36	3	US-09-315-304B-844	Sequence 844, Ap
600	32	47.1	1447	3	US-08-954-668-19	Sequence 19, Appl	673	31	45.6	36	4	US-09-834-784-843	Sequence 843, Ap
601	32	47.1	1447	3	US-09-268-140-5	Sequence 5, Appl	674	31	45.6	36	4	US-09-834-784-844	Sequence 844, Ap
602	32	47.1	1447	4	US-08-918-658-19	Sequence 19, Appl	675	31	45.6	36	4	US-09-515-965A-843	Sequence 843, Ap
603	32	47.1	1447	4	US-09-724-631-19	Sequence 19, Appl	676	31	45.6	36	4	US-09-515-965A-844	Sequence 844, Ap
604	32	47.1	1447	4	US-08-954-701A-19	Sequence 19, Appl	677	31	45.6	36	4	US-09-350-641C-843	Sequence 843, Ap
605	32	47.1	1447	5	PCT-US95-13233-19	Sequence 19, Appl	678	31	45.6	36	4	US-09-350-641C-844	Sequence 844, Ap
606	32	47.1	2616	6	5206163-3	Patent No. 5206163	679	31	45.6	36	4	US-09-350-841A-843	Sequence 843, Ap
607	32	47.1	2616	6	5206163-3	Patent No. 5206163	680	31	45.6	36	4	US-09-350-841A-844	Sequence 844, Ap
608	32	47.1	3170	2	US-07-642-734C-5	Sequence 5, Appl	681	31	45.6	37	3	US-08-230-548-31	Sequence 31, Appl
609	32	47.1	3170	3	US-08-439-009A-5	Sequence 5, Appl	682	31	45.6	40	3	US-08-964-302A-1	Sequence 1, Appl
610	31.5	46.3	130	4	US-09-252-991A-31064	Sequence 31064, A	683	31	45.6	40	3	US-08-964-302A-2	Sequence 2, Appl
611	31.5	46.3	323	4	US-09-902-540-15597	Sequence 15597, A	684	31	45.6	40	3	US-09-441-416A-1	Sequence 1, Appl

31	45.6	40	3	US-09-441-416A-2	Sequence 2, Appli
31	45.6	41	1	US-08-170-852-6	Sequence 6, Appli
31	45.6	41	1	US-08-250-754B-6	Sequence 6, Appli
31	45.6	41	3	US-09-441-416A-21	Sequence 21, Appl
31	45.6	41	3	US-09-441-416A-23	Sequence 23, Appl
31	45.6	41	4	US-09-450-071A-6	Sequence 6, Appli
31	45.6	41	5	PCT-US95-07539-6	Sequence 6, Appli
31	45.6	46	3	US-09-230-54B-25	Sequence 25, Appl
31	45.6	50	1	US-08-464-339A-13	Sequence 16, Appl
31	45.6	51	3	US-09-082-279B-803	Sequence 803, App
31	45.6	51	3	US-09-082-279B-804	Sequence 804, App
31	45.6	51	3	US-09-315-304B-803	Sequence 803, App
31	45.6	51	3	US-09-315-304B-804	Sequence 804, App
31	45.6	51	4	US-09-834-784-803	Sequence 803, App
31	45.6	51	4	US-09-834-784-804	Sequence 804, App
31	45.6	51	4	US-09-515-965A-803	Sequence 803, App
31	45.6	51	4	US-09-515-965A-804	Sequence 804, App
31	45.6	51	4	US-09-350-641C-803	Sequence 803, App
31	45.6	51	4	US-09-350-641C-804	Sequence 804, App
31	45.6	51	4	US-09-350-841A-803	Sequence 803, App
31	45.6	51	4	US-09-350-841A-804	Sequence 804, App
31	45.6	60	4	US-09-248-796A-21719	Sequence 21719, A
31	45.6	68	4	US-09-107-532A-4291	Sequence 4291, Ap
31	45.6	69	4	US-09-149-476-601	Sequence 601, App
31	45.6	70	4	US-09-149-476-605	Sequence 465, App
31	45.6	82	4	US-09-270-767-44099	Sequence 44099, A
31	45.6	94	4	US-09-621-976-7307	Sequence 7307, Ap
31	45.6	96	4	US-09-248-796A-24044	Sequence 24044, A
31	45.6	98	4	US-09-248-796A-18737	Sequence 18737, A
31	45.6	99	4	US-09-540-150-10041	Sequence 10041, A
31	45.6	104	3	US-09-133-521-4	Sequence 4, Appli
31	45.6	104	4	US-09-920-897-2	Sequence 2, Appli
31	45.6	110	4	US-09-732-210-720	Sequence 720, App
31	45.6	111	4	US-09-572-408B-9	Sequence 9, Appli
31	45.6	112	4	US-09-621-976-6045	Sequence 6045, Ap
31	45.6	115	4	US-09-732-210-257	Sequence 257, App
31	45.6	122	4	US-09-513-993C-1516	Sequence 5156, Ap
31	45.6	123	3	US-09-489-039A-10218	Sequence 10218, A
31	45.6	125	4	US-09-902-540-16622	Sequence 16622, A
31	45.6	126	4	US-09-489-039A-8610	Sequence 8610, Ap
31	45.6	127	4	US-09-621-976-3956	Sequence 3956, Ap
31	45.6	128	4	US-09-134-000C-5393	Sequence 5393, Ap
31	45.6	146	4	US-09-902-540-10219	Sequence 10319, A
31	45.6	153	4	US-09-958-413A-6	Sequence 6, Appli
31	45.6	171	4	US-09-489-039A-13877	Sequence 13877, A
31	45.6	173	1	US-07-743-518-26	Sequence 26, Appl
31	45.6	176	2	US-08-484-126-8	Sequence 8, Appli
31	45.6	176	4	US-09-374-909-8	Sequence 8, Appli
31	45.6	178	4	US-09-489-039A-12064	Sequence 12064, A
31	45.6	194	4	US-09-902-540-10702	Sequence 10702, A
31	45.6	205	4	US-09-417-721-8	Sequence 8, Appli
31	45.6	205	6	5175383-6	Patent No. 5175383
31	45.6	205	6	5175383-6	Patent No. 5175383
31	45.6	206	1	US-08-102-691-1	Sequence 1, Appli
31	45.6	206	1	US-08-439-725A-13	Sequence 13, Appl
31	45.6	206	1	US-08-464-590A-15	Sequence 15, Appl
31	45.6	206	1	US-08-441-629-10	Sequence 10, Appl
31	45.6	206	1	US-08-462-169B-12	Sequence 12, Appl
31	45.6	206	2	US-08-207-412B-10	Sequence 10, Appl
31	45.6	206	2	US-08-867-471-13	Sequence 13, Appl
31	45.6				

```

; Sequence 2, Application US/09671089
; Patent No. 6780846
; GENERAL INFORMATION:
; APPLICANT: Q'Mahony, Daniel J.
; APPLICANT: Lambkin, Imelda J.
; TITLE OF INVENTION: MEMBRANE TRANSLOCATING PEPTIDE DRUG DELIVERY SYSTEM
; FILE REFERENCE: E1067/20018
; CURRENT APPLICATION NUMBER: US/09/671,089
; CURRENT FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/156,246
; PRIOR FILING DATE: 1999-09-27
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 2
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: membrane translocating peptide
; FEATURE:
; NAME/KEY: MOD RES
; LOCATION: (15)..(15)
; OTHER INFORMATION: linked to FITC-LC
; US-09-671-089-2

```

## ALIGNMENTS

RESULT 1  
US-09-671-089-2

```
; TITLE OF INVENTION: MEMBRANE TRANSLOCATING PEPTIDE DRUG DELIVERY SYSTEM
; FILE REFERENCE: E1067/20018
; CURRENT APPLICATION NUMBER: US/09/671,089
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/156,246
; PRIOR FILING DATE: 1999-09-27
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 48
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: dansylated membrane translocating peptide
; NAME/KEY: MOD RES
; LOCATION: (1)-(1)
; OTHER INFORMATION: dansylated
US-09-671-089-48

Query Match      100.0%; Score 68; DB 4; Length 16;
Best Local Similarity 100.0%; Pred. No. 6e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKAARAVLLPVLLAAP 15
Db 2 KKAARAVLLPVLLAAP 16

RESULT 4
US-09-671-089-4
; Sequence 4, Application US/09671089
; Patent No. 6780846
; GENERAL INFORMATION:
; APPLICANT: O'Mahony, Daniel J.
; APPLICANT: Lambkin, Inelda J.
; TITLE OF INVENTION: MEMBRANE TRANSLOCATING PEPTIDE DRUG DELIVERY SYSTEM
; FILE REFERENCE: E1067/20018
; CURRENT APPLICATION NUMBER: US/09/671,089
; CURRENT FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/156,246
; PRIOR FILING DATE: 1999-09-27
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: membrane translocating peptide
US-09-671-089-4

Query Match      100.0%; Score 68; DB 4; Length 19;
Best Local Similarity 100.0%; Pred. No. 7.2e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKAARAVLLPVLLAAP 15
Db 1 KKAARAVLLPVLLAAP 15

RESULT 5
US-09-671-089-5
; Sequence 5, Application US/09671089
; Patent No. 6780846
; GENERAL INFORMATION:
; APPLICANT: O'Mahony, Daniel J.
; APPLICANT: Lambkin, Inelda J.
; TITLE OF INVENTION: MEMBRANE TRANSLOCATING PEPTIDE DRUG DELIVERY SYSTEM
; FILE REFERENCE: E1067/20018
; CURRENT APPLICATION NUMBER: US/09/671,089
; CURRENT FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/156,246
```

```
; PRIOR FILING DATE: 1999-09-27
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: membrane translocating peptide, cyclic
US-09-671-089-5

Query Match      94.1%; Score 64; DB 4; Length 16;
Best Local Similarity 93.3%; Pred. No. 0.00027;
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 KKAARAVLLPVLLAAP 15
Db 1 KKAARAVLLPVLLAAP 15

RESULT 6
US-09-186-170-1
; Sequence 1, Application US/09186170
; Patent No. 6248558
; GENERAL INFORMATION:
; APPLICANT: Lin, Yao-Zhong
; APPLICANT: Donahue, John P.
; APPLICANT: Rojas, Mauricio
; APPLICANT: Tan, Zhongjia
; TITLE OF INVENTION: "Sequence and Method for Genetic Engineering of
; Patent No. 6248558
; TITLE OF INVENTION: Proteins with Cell Membrane Translocating Activity"
; FILE REFERENCE: VU9841
; CURRENT APPLICATION NUMBER: US/09/186,170
; CURRENT FILING DATE: 1998-11-04
; EARLIER APPLICATION NUMBER: 60/080,083
; EARLIER FILING DATE: 1998-03-31
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Amino acid
; OTHER INFORMATION: sequence of peptide which transports proteins
; OTHER INFORMATION: through the cell membrane into the cell.
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(12)
; PUBLICATION INFORMATION:
; TITLE: "Genetic Engineering of Proteins with Cell Membrane
; TITLE: "Permeability"
; JOURNAL: Nature Biotechnology
; VOLUME: 16
; ISSUE: April
; PAGES: 370-375
; DATE: 1998-04-01
; RELEVANT RESIDUES: 1 TO 12
US-09-186-170-1

Query Match      79.4%; Score 54; DB 3; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.0089;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 AAVLLPVLLAAP 15
Db 1 AAVLLPVLLAAP 12

RESULT 7
US-09-562-868-1
; Sequence 1, Application US/09562868
```

```
; Patent No. 6432680
; GENERAL INFORMATION:
; APPLICANT: Lin, Yao-Zhong
; APPLICANT: Donahue, John P.
; APPLICANT: Rojas, Mauricio
; APPLICANT: Tan, Zhongjia
; TITLE OF INVENTION: "Sequence and Method for Genetic Engineering of
; Proteins with Cell Membrane Translocating Activity"
; FILE REFERENCE: 22000.009702
; CURRENT APPLICATION NUMBER: US/09/562,868
; CURRENT FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 60/080,083
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 09/186,170
; PRIOR FILING DATE: 1998-11-04
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Amino acid
; sequence of peptide which transports proteins
; OTHER INFORMATION: through the cell membrane into the cell.
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(12)
; PUBLICATION INFORMATION:
; AUTHORS: Rojas, M. et al.
; TITLE: "Genetic Engineering of Proteins with Cell Membrane
; Permeability"
; JOURNAL: Nature Biotechnology
; VOLUME: 16
; ISSUE: April
; PAGES: 370-375
; DATE: 1998-04-01
; US-09-562-868-1

Query Match 79.4%; Score 54; DB 4; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.0089;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 AAVLLPVLLAAP 15
Db 1 AAVLLPVLLAAP 12

RESULT 8
US-09-997-465B-4
; Sequence 4, Application US/09997465B
; Patent No. 6673574
; GENERAL INFORMATION:
; APPLICANT: Stern, William
; APPLICANT: Mehta, No. 6673574er M.
; APPLICANT: Ray, Martha V.L.
; TITLE OF INVENTION: IMPROVED ORAL DELIVERY OF PEPTIDES USING ENZYME-CLEAVABLE MEMBRAN
; TRANSLOCATORS
; FILE REFERENCE: P/546-247
; CURRENT APPLICATION NUMBER: US/09/997,465B
; CURRENT FILING DATE: 2001-11-29
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-997-465B-4

Query Match 79.4%; Score 54; DB 4; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.0089;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 4 AAVLLPVLLAAP 15
Db 1 AAVLLPVLLAAP 12

RESULT 9
US-10-083-889-8
; Sequence 8, Application US/10083889
; Patent No. 6673894
; GENERAL INFORMATION:
; APPLICANT: Zahner, Joseph E.
; TITLE OF INVENTION: Inhibitor of cell proliferation and methods of use thereof.
; FILE REFERENCE: 16850-7331
; CURRENT APPLICATION NUMBER: US/10/083,889
; CURRENT FILING DATE: 2002-02-27
; PRIOR APPLICATION NUMBER: US 60/271,798
; PRIOR FILING DATE: 2001-02-27
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: Microsoft Word 97
; SEQ ID NO 8
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Mammalian
; US-10-083-889-8

Query Match 79.4%; Score 54; DB 4; Length 12;
Best Local Similarity 100.0%; Pred. No. 0.0089;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 AAVLLPVLLAAP 15
Db 1 AAVLLPVLLAAP 12

RESULT 10
US-10-116-288A-1
; Sequence 1, Application US/10116288A
; Patent No. 6780843
; GENERAL INFORMATION:
; APPLICANT: Lin, Yao-Zhong
; APPLICANT: Donahue, John P.
; APPLICANT: Rojas, Mauricio
; APPLICANT: Tan, Zhongjia
; TITLE OF INVENTION: "Sequence and Method for Genetic Engineering of
; Proteins with Cell Membrane Translocating Activity"
; Patent No. 6780843
; FILE REFERENCE: 22000.009703
; CURRENT APPLICATION NUMBER: US/10/116,288A
; CURRENT FILING DATE: 2002-04-04
; PRIOR APPLICATION NUMBER: 09/562,868
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: 09/186,170
; PRIOR FILING DATE: 1998-11-04
; PRIOR APPLICATION NUMBER: 60/080,083
; PRIOR FILING DATE: 1998-03-31
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Amino acid
; sequence of peptide which transports proteins
; OTHER INFORMATION: through the cell membrane into the cell.
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(12)
; PUBLICATION INFORMATION:
; AUTHORS: Rojas, M. et al.
; TITLE: "Genetic Engineering of Proteins with Cell Membrane
; Permeability"
; JOURNAL: Nature Biotechnology
```

VOLUME: 16  
ISSUE: April  
PAGES: 370-375  
DATE: 1998-04-01  
US-10-116-288A-1

Query Match 79.4%; Score 54; DB 4; Length 12;  
Best Local Similarity 100.0%; Pred. No. 0.0089;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 AAVLLPVLLAAP 15  
| | | | | | | | | | | | | |  
DB 1 AAVLLPVLLAAP 12

## RESULT 11

US-09-671-089-1  
; Sequence 1, Application US/09671089  
; Patent No. 6780846  
; GENERAL INFORMATION:  
; APPLICANT: O'Mahony, Daniel J.  
; APPLICANT: Lambkin, Imelda J.  
; TITLE OF INVENTION: MEMBRANE TRANSLOCATING PEPTIDE DRUG DELIVERY SYSTEM  
; FILE REFERENCE: E1067/20018  
; CURRENT APPLICATION NUMBER: US/09/671,089  
; CURRENT FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: 60/156,246  
; PRIOR FILING DATE: 1999-09-27  
; NUMBER OF SEQ ID NOS: 59  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 1  
; LENGTH: 12  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: membrane translocating peptide  
US-09-671-089-1

Query Match 79.4%; Score 54; DB 4; Length 12;  
Best Local Similarity 100.0%; Pred. No. 0.0089;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 AAVLLPVLLAAP 15  
| | | | | | | | | | | | | |  
DB 1 AAVLLPVLLAAP 12

## RESULT 12

US-09-671-089-14  
; Sequence 14, Application US/09671089  
; Patent No. 6780846  
; GENERAL INFORMATION:  
; APPLICANT: O'Mahony, Daniel J.  
; APPLICANT: Lambkin, Imelda J.  
; TITLE OF INVENTION: MEMBRANE TRANSLOCATING PEPTIDE DRUG DELIVERY SYSTEM  
; FILE REFERENCE: E1067/20018  
; CURRENT APPLICATION NUMBER: US/09/671,089  
; CURRENT FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: 60/156,246  
; PRIOR FILING DATE: 1999-09-27  
; NUMBER OF SEQ ID NOS: 59  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 14  
; LENGTH: 12  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: membrane translocating peptide  
US-09-671-089-14

Query Match 79.4%; Score 54; DB 4; Length 12;  
Best Local Similarity 100.0%; Pred. No. 0.0089;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 AAVLLPVLLAAP 15  
| | | | | | | | | | | | | |  
DB 1 AAVLLPVLLAAP 12

## RESULT 13

US-10-144-549-19  
; Sequence 19, Application US/10144549  
; Patent No. 6835810  
; GENERAL INFORMATION:  
; APPLICANT: Geneshuttle Biopharm, Inc.  
; APPLICANT: Hwu, Paul L.  
; TITLE OF INVENTION: A NEW FUSION PROTEIN FOR USE AS VECTOR  
; FILE REFERENCE: MBHB 02-340  
; CURRENT APPLICATION NUMBER: US/10/144,549  
; CURRENT FILING DATE: 2002-05-13  
; NUMBER OF SEQ ID NOS: 31  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 19  
; LENGTH: 12  
; TYPE: PRT  
; ORGANISM: Kaposi's sarcoma-associated herpesvirus  
; FEATURE:  
; NAME/KEY: MISC FEATURE  
; OTHER INFORMATION: Kaposi FGF signal sequence.  
US-10-144-549-19

Query Match 79.4%; Score 54; DB 4; Length 12;  
Best Local Similarity 100.0%; Pred. No. 0.0089;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 AAVLLPVLLAAP 15  
| | | | | | | | | | | | | |  
DB 1 AAVLLPVLLAAP 12

## RESULT 14

US-10-116-288A-22  
; Sequence 22, Application US/10116288A  
; Patent No. 6780843  
; GENERAL INFORMATION:  
; APPLICANT: Lin, Yao-Zhong  
; APPLICANT: Donahue, John P.  
; APPLICANT: Rojas, Mauricio  
; APPLICANT: Tan, Zhongjia  
; TITLE OF INVENTION: Sequence and Method for Genetic Engineering of  
; Patent No. 6780843  
; TITLE OF INVENTION: Proteins with Cell Membrane Translocating Activity"  
; FILE REFERENCE: 22000.0097U3  
; CURRENT APPLICATION NUMBER: US/10/116,288A  
; CURRENT FILING DATE: 2002-04-04  
; PRIOR APPLICATION NUMBER: 09/562,868  
; PRIOR FILING DATE: 2000-05-01  
; PRIOR APPLICATION NUMBER: 09/186,170  
; PRIOR FILING DATE: 1998-11-04  
; PRIOR APPLICATION NUMBER: 60/080,083  
; PRIOR FILING DATE: 1998-03-31  
; NUMBER OF SEQ ID NOS: 23  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 22  
; LENGTH: 19  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Amino acid  
; OTHER INFORMATION: sequence of peptide which transports proteins  
; OTHER INFORMATION: through the cell membrane into the cell.  
US-10-116-288A-22

Query Match 79.4%; Score 54; DB 4; Length 19;  
Best Local Similarity 100.0%; Pred. No. 0.014;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 AAVLLPVLAAAP 15  
|||||  
Db 8 AAVLLPVLAAAP 19

RESULT 15  
US-10-116-288A-20  
; Sequence 20, Application US/10116288A  
; Patent No. 6780843  
; GENERAL INFORMATION:  
; APPLICANT: Lin, Yao-Zhong  
; APPLICANT: Donahue, John P.  
; APPLICANT: Rojase, Mauricio  
; APPLICANT: Tan, Zhongjia  
; TITLE OF INVENTION: "Sequence and Method for Genetic Engineering of  
; Patent No. 6780843  
; TITLE OF INVENTION: Proteins with Cell Membrane Translocating Activity"  
; CURRENT APPLICATION NUMBER: US/10/116,288A  
; CURRENT FILING DATE: 2002-04-04  
; PRIOR APPLICATION NUMBER: 09/562,868  
; PRIOR FILING DATE: 2000-05-01  
; PRIOR APPLICATION NUMBER: 09/186,170  
; PRIOR FILING DATE: 1998-11-04  
; PRIOR APPLICATION NUMBER: 60/080,083  
; PRIOR FILING DATE: 1998-03-31  
; NUMBER OF SEQ ID NOS: 23  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 20  
; LENGTH: 25  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Amino acid  
; OTHER INFORMATION: sequence of peptide which transports proteins  
; OTHER INFORMATION: through the cell membrane into the cell.  
US-10-116-288A-20

Query Match 79.4%; Score 54; DB 4; Length 25;  
Best Local Similarity 100.0%; Pred. No. 0.019;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 AAVLLPVLAAAP 15  
|||||  
Db 7 AAVLLPVLAAAP 18

RESULT 16  
US-09-671-089-7  
; Sequence 7, Application US/09671089  
; Patent No. 6780846  
; GENERAL INFORMATION:  
; APPLICANT: O'Mahony, Daniel J.  
; APPLICANT: Lambkin, Imelda J.  
; TITLE OF INVENTION: MEMBRANE TRANSLOCATING PEPTIDE DRUG DELIVERY SYSTEM  
; FILE REFERENCE: E1067/20018  
; CURRENT APPLICATION NUMBER: US/09/671,089  
; CURRENT FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: 60/156,246  
; PRIOR FILING DATE: 1999-09-27  
; NUMBER OF SEQ ID NOS: 59  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 7  
; LENGTH: 14  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: membrane translocating peptide, cyclic internal  
US-09-671-089-7

Query Match 77.9%; Score 53; DB 4; Length 14;  
Best Local Similarity 92.3%; Pred. No. 0.015;

Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 KKAARAVLLPVLIA 13  
|||||  
Db 1 KKAARAVLLPVLIA 13

RESULT 17  
US-09-671-089-58  
; Sequence 58, Application US/09671089  
; Patent No. 6780846  
; GENERAL INFORMATION:  
; APPLICANT: O'Mahony, Daniel J.  
; APPLICANT: Lambkin, Imelda J.  
; TITLE OF INVENTION: MEMBRANE TRANSLOCATING PEPTIDE DRUG DELIVERY SYSTEM  
; FILE REFERENCE: E1067/20018  
; CURRENT APPLICATION NUMBER: US/09/671,089  
; CURRENT FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: 60/156,246  
; PRIOR FILING DATE: 1999-09-27  
; NUMBER OF SEQ ID NOS: 59  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 58  
; LENGTH: 14  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: membrane translocating peptide, cyclic internal  
US-09-671-089-58

Query Match 77.9%; Score 53; DB 4; Length 14;  
Best Local Similarity 92.3%; Pred. No. 0.015;  
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 KKAARAVLLPVLIA 13  
|||||  
Db 1 KKAARAVLLPVLIA 13

RESULT 18  
US-09-671-089-24  
; Sequence 24, Application US/09671089  
; Patent No. 6780846  
; GENERAL INFORMATION:  
; APPLICANT: O'Mahony, Daniel J.  
; APPLICANT: Lambkin, Imelda J.  
; TITLE OF INVENTION: MEMBRANE TRANSLOCATING PEPTIDE DRUG DELIVERY SYSTEM  
; FILE REFERENCE: E1067/20018  
; CURRENT APPLICATION NUMBER: US/09/671,089  
; CURRENT FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: 60/156,246  
; PRIOR FILING DATE: 1999-09-27  
; NUMBER OF SEQ ID NOS: 59  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 24  
; LENGTH: 17  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: membrane translocating peptide  
US-09-671-089-24

Query Match 75.0%; Score 51; DB 4; Length 17;  
Best Local Similarity 92.3%; Pred. No. 0.04;  
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 KKAARAVLLPVLIA 13  
|||||  
Db 5 KKAARAVLLPVLIA 17

RESULT 19  
US-09-186-170-9

; Sequence 9, Application US/09186170  
; Patent No. 6248558  
; GENERAL INFORMATION:  
; APPLICANT: Lin, Yao-Zhong  
; APPLICANT: Donahue, John P.  
; APPLICANT: Rojas, Mauricio  
; APPLICANT: Tan, Zhongjia  
; TITLE OF INVENTION: "Sequence and Method for Genetic Engineering of  
; Proteins with Cell Membrane Translocating Activity"  
; Patent No. 6248558  
; FILE REFERENCE: VU9841  
; CURRENT APPLICATION NUMBER: US/09/186,170  
; CURRENT FILING DATE: 1998-11-04  
; EARLIER APPLICATION NUMBER: 60/080,083  
; EARLIER FILING DATE: 1998-03-31  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 9  
; LENGTH: 11  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Amino acid  
; OTHER INFORMATION: sequence of peptide which transports proteins  
; OTHER INFORMATION: through the cell membrane into the cell.  
; NAME/KEY: PEPTIDE  
; LOCATION: (1)..(11)  
; TITLE: "Genetic Engineering of Proteins with Cell Membrane  
; TITLE: Permeability"  
; JOURNAL: Nature Biotechnology  
; VOLUME: 16  
; ISSUE: April  
; PAGES: 370-375  
; DATE: 1998-04-01  
; RELEVANT RESIDUES: 1 TO 12  
US-09-186-170-9

Query Match 73.5%; Score 50; DB 3; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.037;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 AVLLPVLLAAP 15  
| | | | | | | | | | | | | | |  
Db 1 AVLLPVLLAAP 11

RESULT 20  
US-09-562-868-9  
; Sequence 9, Application US/09562868  
; Patent No. 6432680  
; GENERAL INFORMATION:  
; APPLICANT: Lin, Yao-Zhong  
; APPLICANT: Donahue, John P.  
; APPLICANT: Rojas, Mauricio  
; APPLICANT: Tan, Zhongjia  
; TITLE OF INVENTION: "Sequence and Method for Genetic Engineering of  
; Proteins with Cell Membrane Translocating Activity"  
; Patent No. 6432680  
; FILE REFERENCE: 22000.009702  
; CURRENT APPLICATION NUMBER: US/09/562,868  
; CURRENT FILING DATE: 2000-05-01  
; PRIOR APPLICATION NUMBER: 60/080,083  
; PRIOR FILING DATE: 1998-03-31  
; PRIOR APPLICATION NUMBER: 09/186,170  
; PRIOR FILING DATE: 1998-11-04  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 9  
; LENGTH: 11  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Amino acid  
; OTHER INFORMATION: sequence of peptide which transports proteins  
; OTHER INFORMATION: through the cell membrane into the cell.  
; NAME/KEY: PEPTIDE  
; LOCATION: (1)..(11)  
; TITLE: "Genetic Engineering of Proteins with Cell Membrane  
; TITLE: Permeability"  
; JOURNAL: Nature Biotechnology  
; VOLUME: 16  
; ISSUE: April  
; PAGES: 370-375  
; DATE: 1998-04-01  
US-10-116-288A-9

Query Match 73.5%; Score 50; DB 4; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.037;

; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Amino acid  
; OTHER INFORMATION: sequence of peptide which transports proteins  
; OTHER INFORMATION: through the cell membrane into the cell.  
; NAME/KEY: PEPTIDE  
; LOCATION: (1)..(11)  
; TITLE: "Genetic Engineering of Proteins with Cell Membrane  
; TITLE: Permeability"  
; JOURNAL: Nature Biotechnology  
; VOLUME: 16  
; ISSUE: April  
; PAGES: 370-375  
; DATE: 1998-04-01  
US-09-562-868-9

Query Match 73.5%; Score 50; DB 4; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.037;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 AVLLPVLLAAP 15  
| | | | | | | | | | | | | | |  
Db 1 AVLLPVLLAAP 11

RESULT 21  
US-10-116-288A-9  
; Sequence 9, Application US/10116288A  
; Patent No. 6780843  
; GENERAL INFORMATION:  
; APPLICANT: Lin, Yao-Zhong  
; APPLICANT: Donahue, John P.  
; APPLICANT: Rojas, Mauricio  
; APPLICANT: Tan, Zhongjia  
; TITLE OF INVENTION: "Sequence and Method for Genetic Engineering of  
; Proteins with Cell Membrane Translocating Activity"  
; Patent No. 6780843  
; FILE REFERENCE: 22000.009703  
; CURRENT APPLICATION NUMBER: US/10/116,288A  
; CURRENT FILING DATE: 2002-04-04  
; PRIOR APPLICATION NUMBER: 09/562,868  
; PRIOR FILING DATE: 2000-05-01  
; PRIOR APPLICATION NUMBER: 09/186,170  
; PRIOR FILING DATE: 1998-11-04  
; PRIOR APPLICATION NUMBER: 60/080,083  
; PRIOR FILING DATE: 1998-03-31  
; NUMBER OF SEQ ID NOS: 23  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 9  
; LENGTH: 11  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Amino acid  
; OTHER INFORMATION: sequence of peptide which transports proteins  
; OTHER INFORMATION: through the cell membrane into the cell.  
; NAME/KEY: PEPTIDE  
; LOCATION: (1)..(11)  
; TITLE: "Genetic Engineering of Proteins with Cell Membrane  
; TITLE: Permeability"  
; JOURNAL: Nature Biotechnology  
; VOLUME: 16  
; ISSUE: April  
; PAGES: 370-375  
; DATE: 1998-04-01  
US-10-116-288A-9

Query Match 73.5%; Score 50; DB 4; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.037;



Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 AVLLPVLLAAP 15  
|||||  
Db 1 AVLLPVLLAAP 11

RESULT 22  
US-09-671-089-19  
; Sequence 19, Application US/09671089  
; Patent No. 6780846  
; GENERAL INFORMATION:  
; APPLICANT: O'Mahony, Daniel J.  
; APPLICANT: Lambkin, Inelda J.  
; TITLE OF INVENTION: MEMBRANE TRANSLOCATING PEPTIDE DRUG DELIVERY SYSTEM  
; FILE REFERENCE: E1067/20018  
; CURRENT APPLICATION NUMBER: US/09/671,089  
; CURRENT FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: 60/156,246  
; PRIOR FILING DATE: 1999-09-27  
; NUMBER OF SEQ ID NOS: 59  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 19  
; LENGTH: 11  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: membrane translocating peptide  
US-09-671-089-19

Query Match 73.5%; Score 50; DB 4; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.037; 0; Indels 0; Gaps 0;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 AVLLPVLLAAP 15  
|||||  
Db 1 AVLLPVLLAAP 11

RESULT 23  
US-09-671-089-10  
; Sequence 10, Application US/09671089  
; Patent No. 6780846  
; GENERAL INFORMATION:  
; APPLICANT: O'Mahony, Daniel J.  
; APPLICANT: Lambkin, Inelda J.  
; TITLE OF INVENTION: MEMBRANE TRANSLOCATING PEPTIDE DRUG DELIVERY SYSTEM  
; FILE REFERENCE: E1067/20018  
; CURRENT APPLICATION NUMBER: US/09/671,089  
; CURRENT FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: 60/156,246  
; PRIOR FILING DATE: 1999-09-27  
; NUMBER OF SEQ ID NOS: 59  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 10  
; LENGTH: 13  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: membrane translocating peptide, cyclic  
US-09-671-089-10

Query Match 73.5%; Score 50; DB 4; Length 13;  
Best Local Similarity 100.0%; Pred. No. 0.044; 0; Indels 0; Gaps 0;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 AVLLPVLLAAP 15  
|||||  
Db 2 AVLLPVLLAAP 12

RESULT 24  
US-09-186-170-5

; Sequence 5, Application US/09186170  
; Patent No. 6248558  
; GENERAL INFORMATION:  
; APPLICANT: Lin, Yao-Zhong  
; APPLICANT: Donahue, John P.  
; APPLICANT: Rojas, Mauricio  
; APPLICANT: Tan, Zhongjia  
; TITLE OF INVENTION: "Sequence and Method for Genetic Engineering of  
; Patent No. 6248558  
; TITLE OF INVENTION: Proteins with Cell Membrane Translocating Activity"  
; FILE REFERENCE: VU9841  
; CURRENT APPLICATION NUMBER: US/09/186,170  
; CURRENT FILING DATE: 1998-11-04  
; EARLIER APPLICATION NUMBER: 60/080,083  
; EARLIER FILING DATE: 1998-03-31  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 5  
; LENGTH: 11  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Amino acid  
; OTHER INFORMATION: sequence of peptide which transports proteins  
; OTHER INFORMATION: through the cell membrane into the cell.  
; FEATURE:  
; NAME/KEY: PEPTIDE  
; LOCATION: (1)..(11)  
; PUBLICATION INFORMATION:  
; TITLE: "Genetic Engineering of Proteins with Cell Membrane  
; TITLE: Permeability"  
; JOURNAL: Nature Biotechnology  
; VOLUME: 16  
; ISSUE: April  
; PAGES: 370-375  
; DATE: 1998-04-01  
; RELEVANT RESIDUES: 1 TO 12  
US-09-186-170-5

Query Match 69.1%; Score 47; DB 3; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.12;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 AAVLLPVLLAA 14  
|||||  
Db 1 AAVLLPVLLAA 11

RESULT 25  
US-09-562-868-5  
; Sequence 5, Application US/09562868  
; Patent No. 6432680  
; GENERAL INFORMATION:  
; APPLICANT: Lin, Yao-Zhong  
; APPLICANT: Donahue, John P.  
; APPLICANT: Rojas, Mauricio  
; APPLICANT: Tan, Zhongjia  
; TITLE OF INVENTION: "Sequence and Method for Genetic Engineering of  
; Patent No. 6432680  
; TITLE OF INVENTION: Proteins with Cell Membrane Translocating Activity"  
; FILE REFERENCE: 22000.0097U2  
; CURRENT APPLICATION NUMBER: US/09/562,868  
; CURRENT FILING DATE: 2000-05-01  
; PRIOR APPLICATION NUMBER: 60/080,083  
; PRIOR FILING DATE: 1998-03-31  
; PRIOR APPLICATION NUMBER: 09/186,170  
; PRIOR FILING DATE: 1998-11-04  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 5  
; LENGTH: 11  
; TYPE: PRT  
; ORGANISM: Artificial Sequence



```
; Sequence 59, Application US/09671089
; Patent No. 6780846
; GENERAL INFORMATION:
; APPLICANT: O'Mahony, Daniel J.
; APPLICANT: Lambkin, Imelda J.
; TITLE OF INVENTION: MEMBRANE TRANSLOCATING PEPTIDE DRUG DELIVERY SYSTEM
; FILE REFERENCE: E1067/20018
; CURRENT APPLICATION NUMBER: US/09/671,089
; CURRENT FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/156,246
; PRIOR FILING DATE: 1999-09-27
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 59
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: membrane translocating peptide
US-09-671-089-59

Query Match          69.1%; Score 47; DB 4; Length 12;
Best Local Similarity 91.7%; Pred. No. 0.13;
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2 KAAAVLLPVLLA 13
DB      1 KAAAVLLPVLLA 12

RESULT 30
US-09-671-089-6
; Sequence 6, Application US/09671089
; Patent No. 6780846
; GENERAL INFORMATION:
; APPLICANT: O'Mahony, Daniel J.
; APPLICANT: Lambkin, Imelda J.
; TITLE OF INVENTION: MEMBRANE TRANSLOCATING PEPTIDE DRUG DELIVERY SYSTEM
; FILE REFERENCE: E1067/20018
; CURRENT APPLICATION NUMBER: US/09/671,089
; CURRENT FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/156,246
; PRIOR FILING DATE: 1999-09-27
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: membrane translocating peptide, cyclic
US-09-671-089-6

Query Match          69.1%; Score 47; DB 4; Length 13;
Best Local Similarity 100.0%; Pred. No. 0.14;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      4 AAVLLPVLLAA 14
DB      2 AAVLLPVLLAA 12

RESULT 31
US-09-671-089-23
; Sequence 23, Application US/09671089
; Patent No. 6780846
; GENERAL INFORMATION:
; APPLICANT: O'Mahony, Daniel J.
; APPLICANT: Lambkin, Imelda J.
; TITLE OF INVENTION: MEMBRANE TRANSLOCATING PEPTIDE DRUG DELIVERY SYSTEM
; FILE REFERENCE: E1067/20018
; CURRENT APPLICATION NUMBER: US/09/671,089
; CURRENT FILING DATE: 2000-09-27
```

```
; PRIOR APPLICATION NUMBER: 60/156,246
; PRIOR FILING DATE: 1999-09-27
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 23
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: membrane translocating peptide
US-09-671-089-23

Query Match          69.1%; Score 47; DB 4; Length 17;
Best Local Similarity 100.0%; Pred. No. 0.18;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      4 AAVLLPVLLAA 14
DB      1 AAVLLPVLLAA 11

RESULT 32
US-09-186-170-8
; Sequence 8, Application US/09186170
; Patent No. 6248558
; GENERAL INFORMATION:
; APPLICANT: Lin, Yao-Zhong
; APPLICANT: Donahue, John P.
; APPLICANT: Rojas, Mauricio
; APPLICANT: Tan, Zhongjia
; TITLE OF INVENTION: "Sequence and Method for Genetic Engineering of
; Patent No. 6248558
; TITLE OF INVENTION: Proteins with Cell Membrane Translocating Activity"
; FILE REFERENCE: VU9841
; CURRENT APPLICATION NUMBER: US/09/186,170
; CURRENT FILING DATE: 1998-11-04
; EARLIER APPLICATION NUMBER: 60/080,083
; EARLIER FILING DATE: 1998-03-31
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Amino acid
; OTHER INFORMATION: sequence of peptide which transports proteins
; OTHER INFORMATION: through the cell membrane into the cell.
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(10)
; PUBLICATION INFORMATION:
; TITLE: "Genetic Engineering of Proteins with Cell Membrane
; TITLE: "Permeability"
; JOURNAL: Nature Biotechnology
; VOLUME: 16
; ISSUE: April
; PAGES: 370-375
; DATE: 1998-04-01
; RELEVANT RESIDUES: 1 TO 12
US-09-186-170-8

Query Match          67.6%; Score 46; DB 3; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.15;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      6 VLLPVLLAAP 15
DB      1 VLLPVLLAAP 10

RESULT 33
US-09-562-868-8
```

; Sequence 8, Application US/09562868  
; Patent No. 6432680  
; GENERAL INFORMATION:  
; APPLICANT: Lin, Yao-Zhong  
; APPLICANT: Donahue, John P.  
; APPLICANT: Rojas, Mauricio  
; APPLICANT: Tan, Zhongjia  
; TITLE OF INVENTION: "Sequence and Method for Genetic Engineering of  
; Proteins with Cell Membrane Translocating Activity"  
; Patent No. 6432680  
; TITLE OF INVENTION: Proteins with Cell Membrane Translocating Activity"  
; FILE REFERENCE: 22000.009702  
; CURRENT APPLICATION NUMBER: US/09/562,868  
; CURRENT FILING DATE: 2000-05-01  
; PRIOR APPLICATION NUMBER: 60/080,083  
; PRIOR FILING DATE: 1998-03-31  
; PRIOR APPLICATION NUMBER: 09/186,170  
; PRIOR FILING DATE: 1998-11-04  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 8  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Amino acid  
; sequence of peptide which transports proteins  
; OTHER INFORMATION: through the cell membrane into the cell.  
; NAME/KEY: PEPTIDE  
; LOCATION: (1)..(10)  
; AUTHORS: Rojas, M. et al.  
; TITLE: "Genetic Engineering of Proteins with Cell Membrane  
; TITLE: Permeability"  
; JOURNAL: Nature Biotechnology  
; VOLUME: 16  
; ISSUE: April  
; PAGES: 370-375  
; DATE: 1998-04-01  
; US-09-562-868-8

Query Match 67.6%; Score 46; DB 4; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.15;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 6 VLLPVLLAAP 15  
| | | | | | | | | |  
Db 1 VLLPVLLAAP 10

RESULT 34  
US-10-116-288A-8  
; Sequence 8, Application US/10116288A  
; Patent No. 6780843  
; GENERAL INFORMATION:  
; APPLICANT: Lin, Yao-Zhong  
; APPLICANT: Donahue, John P.  
; APPLICANT: Rojas, Mauricio  
; APPLICANT: Tan, Zhongjia  
; TITLE OF INVENTION: "Sequence and Method for Genetic Engineering of  
; Proteins with Cell Membrane Translocating Activity"  
; Patent No. 6780843  
; TITLE OF INVENTION: Proteins with Cell Membrane Translocating Activity"  
; FILE REFERENCE: 22000.009703  
; CURRENT APPLICATION NUMBER: US/10/116,288A  
; CURRENT FILING DATE: 2002-04-04  
; PRIOR APPLICATION NUMBER: 09/562,868  
; PRIOR FILING DATE: 2000-05-01  
; PRIOR APPLICATION NUMBER: 09/186,170  
; PRIOR FILING DATE: 1998-11-04  
; PRIOR APPLICATION NUMBER: 60/080,083  
; PRIOR FILING DATE: 1998-03-31  
; NUMBER OF SEQ ID NOS: 23  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 8

; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Amino acid  
; sequence of peptide which transports proteins  
; OTHER INFORMATION: through the cell membrane into the cell.  
; NAME/KEY: PEPTIDE  
; LOCATION: (1)..(10)  
; PUBLICATION INFORMATION:  
; AUTHORS: Rojas, M. et al.  
; TITLE: "Genetic Engineering of Proteins with Cell Membrane  
; TITLE: Permeability"  
; JOURNAL: Nature Biotechnology  
; VOLUME: 16  
; ISSUE: April  
; PAGES: 370-375  
; DATE: 1998-04-01  
; US-10-116-288A-8

Query Match 67.6%; Score 46; DB 4; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.15;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 VLLPVLLAAP 15  
| | | | | | | | | |  
Db 1 VLLPVLLAAP 10

RESULT 35  
US-09-671-089-20  
; Sequence 20, Application US/09671089  
; Patent No. 6780846  
; GENERAL INFORMATION:  
; APPLICANT: O'Mahony, Daniel J.  
; APPLICANT: Lambkin, Imelda J.  
; TITLE OF INVENTION: MEMBRANE TRANSLOCATING PEPTIDE DRUG DELIVERY SYSTEM  
; FILE REFERENCE: E1067/20018  
; CURRENT APPLICATION NUMBER: US/09/671,089  
; CURRENT FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: 60/156,246  
; PRIOR FILING DATE: 1999-09-27  
; NUMBER OF SEQ ID NOS: 59  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 20  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: membrane translocating peptide  
; US-09-671-089-20

Query Match 67.6%; Score 46; DB 4; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.15;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 VLLPVLLAAP 15  
| | | | | | | | | |  
Db 1 VLLPVLLAAP 10

RESULT 36  
US-09-671-089-11  
; Sequence 11, Application US/09671089  
; Patent No. 6780846  
; GENERAL INFORMATION:  
; APPLICANT: O'Mahony, Daniel J.  
; APPLICANT: Lambkin, Imelda J.  
; TITLE OF INVENTION: MEMBRANE TRANSLOCATING PEPTIDE DRUG DELIVERY SYSTEM  
; FILE REFERENCE: E1067/20018  
; CURRENT APPLICATION NUMBER: US/09/671,089  
; CURRENT FILING DATE: 2000-09-27

```
; SOFTWARE: PATEMILLI VER. 2.0  
; SEQ ID NO 4
```

; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Amino acid  
; OTHER INFORMATION: sequence of peptide which transports proteins  
; OTHER INFORMATION: through the cell membrane into the cell.  
; FEATURE:  
; NAME/KEY: PEPTIDE  
; LOCATION: (1)..(10)  
; PUBLICATION INFORMATION:  
; AUTHORS: Rojas, M. et al.  
; TITLE: "Genetic Engineering of Proteins with Cell Membrane  
; TITLE: Permeability"  
; JOURNAL: Nature Biotechnology  
; VOLUME: 16  
; ISSUE: April  
; PAGES: 370-375  
; DATE: 1998-04-01  
US-10-116-288A-4

Query Match 63.2%; Score 43; DB 4; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.47; Indels 0; Gaps 0;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 AAVLLPVLLA 13  
| | | | | | | | | |  
Db 1 AAVLLPVLLA 10

## RESULT 40

US-09-543-681A-8048  
; Sequence 8048, Application US/09543681A  
; Patent No. 6605709  
; GENERAL INFORMATION:  
; APPLICANT: GARY BRETON

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS  
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 2709.1002-001  
; CURRENT APPLICATION NUMBER: US/09/543,681A  
; CURRENT FILING DATE: 2000-04-05  
; PRIOR APPLICATION NUMBER: US 60/128,706  
; PRIOR FILING DATE: 1999-04-09  
; NUMBER OF SEQ ID NOS: 8344  
; SEQ ID NO 8048  
; LENGTH: 191  
; TYPE: PRT  
; ORGANISM: Proteus mirabilis  
US-09-543-681A-8048

Query Match 63.2%; Score 43; DB 4; Length 191;  
Best Local Similarity 72.7%; Pred. No. 10; Indels 0; Gaps 0;  
Matches 8; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KKAARVLLPVL 11  
| | | | | | | | | |  
Db 31 KKAARVLLPIL 41

## RESULT 41

US-09-186-170-7  
; Sequence 7, Application US/09186170  
; Patent No. 6248558  
; GENERAL INFORMATION:  
; APPLICANT: Lin, Yao-Zhong

; APPLICANT: Donahue, John P.  
; APPLICANT: Rojas, Mauricio  
; APPLICANT: Tan, Zhongjia  
; TITLE OF INVENTION: "Sequence and Method for Genetic Engineering of  
; Patent No. 6248558  
; TITLE OF INVENTION: Proteins with Cell Membrane Translocating Activity"  
; FILE REFERENCE: VU9841  
; CURRENT APPLICATION NUMBER: US/09/186,170

; CURRENT FILING DATE: 1998-11-04  
; EARLIER APPLICATION NUMBER: 60/080,083  
; EARLIER FILING DATE: 1998-03-31  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 7  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Amino acid  
; OTHER INFORMATION: sequence of peptide which transports proteins  
; OTHER INFORMATION: through the cell membrane into the cell.  
; FEATURE:  
; NAME/KEY: PEPTIDE  
; LOCATION: (1)..(9)  
; PUBLICATION INFORMATION:  
; TITLE: "Genetic Engineering of Proteins with Cell Membrane  
; TITLE: Permeability"  
; JOURNAL: Nature Biotechnology  
; VOLUME: 16  
; ISSUE: April  
; PAGES: 370-375  
; DATE: 1998-04-01  
; RELEVANT RESIDUES: 1 TO 12  
US-09-186-170-7

Query Match 61.8%; Score 42; DB 3; Length 9;  
Best Local Similarity 100.0%; Pred. No. 4.1e+05; Indels 0; Gaps 0;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 LLPVLLAAP 15  
| | | | | | | | | |  
Db 1 LLPVLLAAP 9

## RESULT 42

US-09-562-868-7  
; Sequence 7, Application US/09562868  
; Patent No. 6432680  
; GENERAL INFORMATION:  
; APPLICANT: Lin, Yao-Zhong

; APPLICANT: Donahue, John P.  
; APPLICANT: Rojas, Mauricio  
; APPLICANT: Tan, Zhongjia  
; TITLE OF INVENTION: "Sequence and Method for Genetic Engineering of  
; Patent No. 6432680  
; TITLE OF INVENTION: Proteins with Cell Membrane Translocating Activity"  
; FILE REFERENCE: 22000.0097U2  
; CURRENT APPLICATION NUMBER: US/09/562,868  
; CURRENT FILING DATE: 2000-05-01  
; PRIOR APPLICATION NUMBER: 60/080,083  
; PRIOR FILING DATE: 1998-03-31  
; PRIOR APPLICATION NUMBER: 09/186,170  
; PRIOR FILING DATE: 1998-11-04  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 7  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Amino acid  
; OTHER INFORMATION: sequence of peptide which transports proteins  
; OTHER INFORMATION: through the cell membrane into the cell.  
; NAME/KEY: PEPTIDE  
; LOCATION: (1)..(9)  
; PUBLICATION INFORMATION:  
; AUTHORS: Rojas, M. et al.  
; TITLE: "Genetic Engineering of Proteins with Cell Membrane  
; TITLE: Permeability"  
; JOURNAL: Nature Biotechnology  
; VOLUME: 16

ISSUE: April  
PAGES: 370-375  
DATE: 1998-04-01  
US-09-562-868-7

Query Match 61.8%; Score 42; DB 4; Length 9;  
Best Local Similarity 100.0%; Pred. No. 4.1e+05;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 LLPVLLAAP 15  
| | | | | | | | | |  
Db 1 LLPVLLAAP 9

RESULT 43  
US-10-116-288A-7  
; Sequence 7, Application US/10116288A  
; Patent No. 6780843  
; GENERAL INFORMATION:  
; APPLICANT: Lin, Yao-Zhong  
; APPLICANT: Donahue, John P.  
; APPLICANT: Rojas, Mauricio  
; APPLICANT: Tan, Zhongjia  
; TITLE OF INVENTION: "Sequence and Method for Genetic Engineering of  
; Patent No. 6780843  
; TITLE OF INVENTION: Proteins with Cell Membrane Translocating Activity"  
; FILE REFERENCE: 22000.0097U3  
; CURRENT APPLICATION NUMBER: US/10/116,288A  
; CURRENT FILING DATE: 2002-04-04  
; PRIOR APPLICATION NUMBER: 09/562,868  
; PRIOR FILING DATE: 2000-05-01  
; PRIOR APPLICATION NUMBER: 09/186,170  
; PRIOR FILING DATE: 1998-11-04  
; PRIOR APPLICATION NUMBER: 60/080,083  
; PRIOR FILING DATE: 1998-03-31  
; NUMBER OF SEQ ID NOS: 23  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 7  
; LENGTH: 9  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Amino acid  
; OTHER INFORMATION: sequence of peptide which transports proteins  
; OTHER INFORMATION: through the cell membrane into the cell.

NAME/KEY: PEPTIDE  
LOCATION: (1)..(9)  
PUBLICATION INFORMATION:  
AUTHORS: Rojas, M. et al.  
TITLE: "Genetic Engineering of Proteins with Cell Membrane  
TITLE: Permeability"  
JOURNAL: Nature Biotechnology  
VOLUME: 16  
ISSUE: April  
PAGES: 370-375  
DATE: 1998-04-01  
US-10-116-288A-7

Query Match 61.8%; Score 42; DB 4; Length 9;  
Best Local Similarity 100.0%; Pred. No. 4.1e+05;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 LLPVLLAAP 15  
| | | | | | | | | |  
Db 1 LLPVLLAAP 9

RESULT 44  
US-09-671-089-21  
; Sequence 21, Application US/09671089  
; Patent No. 6780846  
; GENERAL INFORMATION:

APPLICANT: O'Mahony, Daniel J.  
APPLICANT: Lambkin, Imelda J.  
TITLE OF INVENTION: MEMBRANE TRANSLOCATING PEPTIDE DRUG DELIVERY SYSTEM  
FILE REFERENCE: E1067/20018  
CURRENT APPLICATION NUMBER: US/09/671,089  
CURRENT FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: 60/156,246  
PRIOR FILING DATE: 1999-09-27  
NUMBER OF SEQ ID NOS: 59  
SOFTWARE: Patent in version 3.1  
SEQ ID NO 21  
LENGTH: 9  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: membrane translocating peptide  
US-09-671-089-21

Query Match 61.8%; Score 42; DB 4; Length 9;  
Best Local Similarity 100.0%; Pred. No. 4.1e+05;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 LLPVLLAAP 15  
| | | | | | | | | |  
Db 1 LLPVLLAAP 9

RESULT 45  
US-09-671-089-12  
; Sequence 12, Application US/09671089  
; Patent No. 6780846  
; GENERAL INFORMATION:  
; APPLICANT: O'Mahony, Daniel J.  
; APPLICANT: Lambkin, Imelda J.  
; TITLE OF INVENTION: MEMBRANE TRANSLOCATING PEPTIDE DRUG DELIVERY SYSTEM  
; FILE REFERENCE: E1067/20018  
; CURRENT APPLICATION NUMBER: US/09/671,089  
; CURRENT FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: 60/156,246  
; PRIOR FILING DATE: 1999-09-27  
; NUMBER OF SEQ ID NOS: 59  
; SOFTWARE: Patent In version 3.1  
; SEQ ID NO 12  
; LENGTH: 11  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: membrane translocating peptide, cyclic  
US-09-671-089-12

Query Match 61.8%; Score 42; DB 4; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.76;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 LLPVLLAAP 15  
| | | | | | | | | |  
Db 2 LLPVLLAAP 10

RESULT 46  
US-09-252-991A-18288  
; Sequence 18288, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190



; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 18288  
; LENGTH: 578  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-18288

Query Match 60.3%; Score 41; DB 4; Length 578;  
Best Local Similarity 57.1%; Pred. No. 68;  
Matches 8; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KKAVALLPVLLA 14  
Db 182 RRLAVLLPILL 195

RESULT 47  
US-08-448-489-7  
; Sequence 7, Application US/08448489  
; Patent No. 6184022  
; GENERAL INFORMATION:  
; APPLICANT: SEIKI, Motoharu  
; APPLICANT: SATO, Hiroshi  
; APPLICANT: SHINAGAWA, Akira  
; TITLE OF INVENTION: NOVEL METALLOPROTEINASE AND ENCODING DNA THEREFOR  
; FILE REFERENCE: 55-290P  
; CURRENT APPLICATION NUMBER: US/08/448,489  
; CURRENT FILING DATE: 1995-06-07  
; NUMBER OF SEQ ID NOS: 19  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 7  
; LENGTH: 30  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-08-448-489-7

Query Match 58.8%; Score 40; DB 3; Length 30;  
Best Local Similarity 90.0%; Pred. No. 4.6;  
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 3 AAVALPVL 12  
Db 7 AAVALPVL 16

RESULT 48  
US-09-689-730-7  
; Sequence 7, Application US/09689730  
; Patent No. 6825024  
; GENERAL INFORMATION:  
; APPLICANT: SEIKI, Motoharu  
; APPLICANT: SATO, Hiroshi  
; APPLICANT: SHINAGAWA, Akira  
; TITLE OF INVENTION: NOVEL METALLOPROTEINASE AND ENCODING DNA THEREFOR  
; FILE REFERENCE: 55-290P  
; CURRENT APPLICATION NUMBER: US/09/689,730  
; CURRENT FILING DATE: 2000-10-13  
; PRIOR APPLICATION NUMBER: US/08/448,489  
; PRIOR FILING DATE: 1995-06-07  
; NUMBER OF SEQ ID NOS: 19  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 7  
; LENGTH: 30  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-689-730-7

Query Match 58.8%; Score 40; DB 4; Length 30;  
Best Local Similarity 90.0%; Pred. No. 4.6;  
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 3 AAVALPVL 12

Db 7 AAVALPVL 16

RESULT 49  
US-09-902-540-12703  
; Sequence 12703, Application US/09902540  
; Patent No. 6833447  
; GENERAL INFORMATION:  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Wiegand, Roger C.  
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof  
; FILE REFERENCE: 38-10(15849)B  
; CURRENT APPLICATION NUMBER: US/09/902,540  
; CURRENT FILING DATE: 2001-07-10  
; PRIOR APPLICATION NUMBER: 60/217,883  
; PRIOR FILING DATE: 2000-07-10  
; NUMBER OF SEQ ID NOS: 16825  
; SEQ ID NO 12703  
; LENGTH: 263  
; TYPE: PRT  
; ORGANISM: Myxococcus xanthus  
US-09-902-540-12703

Query Match 58.8%; Score 40; DB 4; Length 263;  
Best Local Similarity 81.8%; Pred. No. 44;  
Matches 9; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 KAAVALPVL 12  
Db 196 KAAVALPVL 206

RESULT 50  
US-09-252-991A-20502  
; Sequence 20502, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 20502  
; LENGTH: 456  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-20502

Query Match 58.8%; Score 40; DB 4; Length 456;  
Best Local Similarity 60.0%; Pred. No. 78;  
Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 1 KKAVALPVL 15  
Db 20 ERAGAVACPVL 34

Search completed: June 2, 2005, 01:40:38  
Job time : 34.125 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 2, 2005, 01:26:16 ; Search time 11.7188 Seconds

(without alignments)  
123.157 Million cell updates/sec

Title: US-10-764-235-2

Perfect score: 68

Sequence: 1 KKAAYLLPVLLAAP 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 750 summaries

Database :

PIR\_79:\*

1: pir1:\*

2: pir2:\*

3: pir3:\*

4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	42	61.8	297	2 AF3275	transporter, dme f
2	42	61.8	939	2 T05209	hypothetical prote
3	41	60.3	1172	2 D72482	probable antibiotic
4	40	58.8	117	2 T09155	lipid transfer pro
5	40	58.8	158	2 A12821	conserved hypothet
6	40	58.8	163	2 A97600	hypothetical prote
7	40	58.8	390	2 A72698	hypothetical prote
8	40	58.8	431	2 AF2635	exopolysaccharide
9	40	58.8	435	2 E97417	exoQ-like protein
10	40	58.8	582	2 I48673	matrix metalloprot
11	40	58.8	582	2 I38028	matrix metalloprot
12	40	58.8	582	2 I8471	matrix metalloprot
13	40	58.8	606	2 E90536	hypothetical prote
14	40	58.8	1247	1 MWIUND	nidogen precursor
15	39	57.4	175	2 B44102	di-N-acetylchitobi
16	39	57.4	201	2 E84187	hypothetical prote
17	39	57.4	227	2 T47186	hypothetical prote
18	39	57.4	252	2 H72469	hypothetical prote
19	39	57.4	303	2 B83336	hypothetical prote
20	39	57.4	354	2 AC0389	phosphate binding
21	39	57.4	382	2 G84971	ompF-like porin (i
22	39	57.4	389	2 D72554	hypothetical prote
23	39	57.4	432	2 T35263	probable carboxype
24	39	57.4	493	2 AB3032	hypothetical prote
25	39	57.4	534	2 F90031	PTS system, arbiti
26	39	57.4	548	2 T51035	hypothetical prote
27	39	57.4	602	2 A96254	exsE protein (Au22
28	38	55.9	155	1 T12M1	trypsin/factor XII
29	38	55.9	182	2 S46667	meiosis-specific p

hypothetical prote  
probable secreted  
probable chitinase  
chitinase (EC 3.2.  
outer membrane por  
outer membrane por  
chitinase (EC 3.2.  
outer membrane por  
alkaline proteinase  
riboflavin biosynt  
probable regulator  
probable ABC-type  
probable membrane  
probable membrane  
hypothetical prote  
probable cation-tr  
cytochrome c3 prec  
cytochrome c-type  
conserved hypothet  
hypothetical prote  
hypothetical prote  
hypothetical prote  
hypothetical prote  
hypothetical prote  
probable fusA prot  
helicase, probable  
exonuclease V (imp  
hypothetical prote  
probable lipoprote  
complement compone  
hypothetical prote  
hypothetical prote  
hypothetical prote  
hypothetical prote  
hypothetical prote  
nodulin-26-like pr  
conserved hypothet  
hypothetical prote  
hypothetical prote  
hypothetical prote  
nodulin-26-like pr  
probable conjugal  
hypothetical prote  
probable integral  
hypothetical prote  
Xaa-Pro dipeptidase  
hypothetical prote  
hypothetical prote  
proteinase (import  
probable membrane  
3-phytase, fusion,  
collagen alpha 1(V  
protein kinase - f  
dynein heavy chain  
hypothetical prote  
genome polypeptide  
cuticle protein lm  
hypothetical prote  
hydrophobin C - cu  
hypothetical prote  
thymic shared anti  
hypothetical prote  
hypothetical prote  
heat shock protein  
aurin precursor (i  
hypothetical prote  
outer membrane por  
hypothetical prote  
conserved hypothet

103	35	51.5	206	2	JC4268	fibroblast growth	176	34	50.0	212	2	PN0565	iduronate-2-sulfat
104	35	51.5	206	2	JC7320	K562 cell-derived	177	34	50.0	217	2	G70407	hypothetical prote
105	35	51.5	214	2	F90966	probable outer mem	178	34	50.0	245	2	G81032	DNA polymerase, ba
106	35	51.5	224	2	T34686	probable integral	179	34	50.0	245	2	D81977	hypothetical prote
107	35	51.5	224	2	F85814	probable outer mem	180	34	50.0	257	2	G83481	conserved hypothet
108	35	51.5	234	2	C83587	hypothetical prote	181	34	50.0	260	1	F64149	nitrate transport
109	35	51.5	238	2	AD0109	thiol, disulfide in	182	34	50.0	275	2	H82929	ATP synthase A cha
110	35	51.5	263	2	S78364	conserved hypothet	183	34	50.0	281	2	C95932	probable sugar upt
111	35	51.5	267	2	D97565	transport protein	184	34	50.0	281	2	D70682	hypothetical prote
112	35	51.5	267	2	AD2786	SEC-independent pr	185	34	50.0	284	2	S52003	major intrinsic pr
113	35	51.5	280	2	AC0411	probable priplasm	186	34	50.0	290	2	E86284	F9i1.2 protein - A
114	35	51.5	296	2	B39581	GREAS protein 9u	187	34	50.0	297	2	D83825	hypothetical prote
115	35	51.5	302	2	AF2651	ABC transporter, m	188	34	50.0	297	2	F87567	integral membrane
116	35	51.5	310	2	T26710	hypothetical prote	189	34	50.0	300	2	D87497	MtCB family protei
117	35	51.5	320	2	A39581	ESAG protein 9c -	190	34	50.0	306	2	AF1142	ABC transporter (b
118	35	51.5	321	2	E97433	polyamine transpor	191	34	50.0	306	2	AI1500	ABC transporter (b
119	35	51.5	328	2	T36494	probable membrane	192	34	50.0	307	1	PNBSL	beta-lactamase (EC
120	35	51.5	351	2	T08015	probable malate de	193	34	50.0	314	2	F75504	probable pili, ty
121	35	51.5	351	2	JC7280	cytokine receptor-	194	34	50.0	323	2	A58892	NADH2 dehydrogenas
122	35	51.5	363	2	AE0616	outer membrane pro	195	34	50.0	328	2	T30999	hypothetical prote
123	35	51.5	363	2	S43159	outer membrane por	196	34	50.0	335	2	T44728	probable hydrolase
124	35	51.5	376	2	T30277	hypothetical prote	197	34	50.0	360	2	AE0172	probable outer mem
125	35	51.5	376	2	T30979	hypothetical prote	198	34	50.0	363	2	G83954	flagellin-associate
126	35	51.5	394	2	AC0755	outer membrane pro	199	34	50.0	368	2	D87271	general secretion
127	35	51.5	394	2	S34263	outer membrane por	200	34	50.0	387	2	F84392	hypothetical prote
128	35	51.5	401	2	D83760	hypothetical prote	201	34	50.0	388	2	A72497	hypothetical prote
129	35	51.5	409	1	S32905	serine proteinase	202	34	50.0	388	2	T26360	hypothetical prote
130	35	51.5	418	2	G87581	conjugal transfer	203	34	50.0	396	2	E83188	1-deoxy-d-xylulose
131	35	51.5	423	2	AI2189	hypothetical prote	204	34	50.0	396	2	B83461	hypothetical prote
132	35	51.5	500	2	E72687	probable branched-	205	34	50.0	404	2	T36239	probable integral
133	35	51.5	514	2	B86164	hypothetical prote	206	34	50.0	412	2	E98253	hypothetical prote
134	35	51.5	533	2	AG1018	probable exported	207	34	50.0	416	2	D75531	folyl-polyglutamat
135	35	51.5	535	2	S78598	D-ribulokinase (EC	208	34	50.0	417	2	A36965	malonyl-CoA decarb
136	35	51.5	549	2	G83201	conserved hypothet	209	34	50.0	434	2	G98235	hypothetical prote
137	35	51.5	554	2	G82538	conserved hypothet	210	34	50.0	439	2	T25494	probable cytochrom
138	35	51.5	589	2	T00698	probable mandeloni	211	34	50.0	445	2	AE3050	two component sens
139	35	51.5	616	2	T00894	hypothetical prote	212	34	50.0	458	2	D96928	PTS system, fructo
140	35	51.5	638	2	A36929	virulence regulato	213	34	50.0	464	2	B64173	hypothetical prote
141	35	51.5	643	2	D86167	protein F21B7.27 [	214	34	50.0	467	2	T34617	NADH2 dehydrogenas
142	35	51.5	677	2	S65573	phosphate-specific	215	34	50.0	497	2	AG2724	ABC transporter, s
143	35	51.5	741	2	T31164	hypothetical prote	216	34	50.0	504	2	T16526	hypothetical prote
144	35	51.5	825	2	E73508	conserved hypothet	217	34	50.0	509	2	C97506	probable solute-bi
145	35	51.5	842	2	S18462	glycoprotein H pre	218	34	50.0	529	1	VGNVAC	major envelope gly
146	35	51.5	863	2	T38016	importin beta-1 su	219	34	50.0	537	2	S67434	probable sterol O-
147	35	51.5	918	2	E90542	lipoprotein (impor	220	34	50.0	557	2	T28150	probable ATP-bind
148	35	51.5	980	2	H84632	probable receptor-	221	34	50.0	563	2	A47153	iduronate-2-sulfat
149	35	51.5	1333	2	A37488	Ras guanine nucleo	222	34	50.0	576	2	S50113	collin, p80 - huma
150	35	51.5	1336	2	S25716	Ras guanine nucleo	223	34	50.0	589	2	D64125	hypothetical prote
151	35	51.5	1581	1	VGM7BV	peplomeric glycoprot	224	34	50.0	592	2	I49239	vesicle transport
152	35	51.5	2206	2	JC5280	voltage-dependent	225	34	50.0	606	2	D95974	probable ABC trans
153	34.5	50.7	308	2	F87283	methionyl-tRNA for	226	34	50.0	608	2	T40814	probable nucleolar
154	34	50.0	37	2	T11961	hypothetical prote	227	34	50.0	610	2	T06280	probable starch sy
155	34	50.0	62	2	T30977	hypothetical prote	228	34	50.0	617	2	E64734	yach protein - Esc
156	34	50.0	74	2	D84231	hypothetical prote	229	34	50.0	622	2	T01414	starch synthase (E
157	34	50.0	85	2	A75190	hypothetical prote	230	34	50.0	625	2	T10661	serine/threonine-s
158	34	50.0	89	2	A63833	hypothetical prote	231	34	50.0	626	2	JQ2322	starch synthase (E
159	34	50.0	91	2	G69750	hypothetical prote	232	34	50.0	635	2	T46407	probable RNA helic
160	34	50.0	94	2	A69105	hypothetical prote	233	34	50.0	637	2	I49638	probable RNA helic
161	34	50.0	104	2	F72753	hypothetical prote	234	34	50.0	660	2	A32378	probable ATP-depen
162	34	50.0	111	2	C69168	conserved hypothet	235	34	50.0	661	2	A58768	ATP-dependent RNA
163	34	50.0	120	2	H95330	hypothetical prote	236	34	50.0	662	1	I84741	RNA helicase - mou
164	34	50.0	123	2	G90878	hypothetical prote	237	34	50.0	669	2	T43866	phosphate transpor
165	34	50.0	123	2	C85740	partial probable o	238	34	50.0	697	1	S13654	ATP-dependent RNA
166	34	50.0	127	2	B69767	conserved hypothet	239	34	50.0	700	2	I51235	DEAD box protein -
167	34	50.0	129	2	H87659	conserved hypothet	240	34	50.0	701	2	S31150	translation elonga
168	34	50.0	135	2	AF0184	probable membrane	241	34	50.0	701	2	H87143	elongation factor
169	34	50.0	142	2	C75271	hypothetical prote	242	34	50.0	703	2	A49994	sulfate transporte
170	34	50.0	168	2	S16031	alpha-amylase inhi	243	34	50.0	713	2	JC2534	RVLG protein - rat
171	34	50.0	168	2	S10029	alpha-amylase inhi	244	34	50.0	798	2	AE1263	exodeoxyribonuclea
172	34	50.0	183	2	F75086	hypothetical prote	245	34	50.0	798	2	AG1625	exodeoxyribonuclea
173	34	50.0	198	2	D87484	glutamine amidotra	246	34	50.0	812	2	D49664	transcription-repa
174	34	50.0	210	2	AB2853	NTP pyrophosphohyd	247	34	50.0	834	2	D86842	hypothetical prote
175	34	50.0	210	2	H97629	mutT/nudix family	248	34	50.0	921	2	S35160	aspartate kinase (

249	34	50.0	977	2	IS2657	seizure-related pr	322	33	48.5	423	2	H81200	cell division prot
250	34	50.0	995	2	C84145	hypothetical prote	323	33	48.5	427	2	T36167	hypothetical prote
251	34	50.0	1089	2	S53978	PSGI protein - yea	324	33	48.5	430	2	T45437	hypothetical prote
252	34	50.0	1093	2	B86748	hypothetical prote	325	33	48.5	435	2	C81777	cell division prot
253	34	50.0	1257	2	AC2841	methionine synthas	326	33	48.5	437	2	T35519	hypothetical prote
254	34	50.0	1259	2	E97618	methionine synthas	327	33	48.5	451	2	A36718	hypothetical prote
255	34	50.0	1261	2	A13471	5-methyltetrahydro	328	33	48.5	451	2	S20723	carboxypeptidase (
256	34	50.0	1353	2	J00407	xanthine dehydroge	329	33	48.5	458	2	H82200	sodium/alanine sym
257	34	50.0	1456	2	T15961	hypothetical prote	330	33	48.5	460	2	A11897	muscarinic acetylch
258	34	50.0	2672	2	A48126	translation activa	331	33	48.5	460	2	I51837	muscarinic acetylch
259	33.5	49.3	78	2	D82110	conserved hypothet	332	33	48.5	460	2	S09508	muscarinic acetylch
260	33.5	49.3	356	2	S34649	genome polyprotein	333	33	48.5	460	2	A24325	muscarinic acetylch
261	33	48.5	42	2	A37907	thyrotropin beta c	334	33	48.5	461	2	A29514	muscarinic acetylch
262	33	48.5	102	2	D83713	hypothetical prote	335	33	48.5	461	2	B10800	muscarinic acetylch
263	33	48.5	119	1	A46492	beta-2-microglobul	336	33	48.5	461	2	E81862	probable integral tr
264	33	48.5	128	2	C83161	hypothetical prote	337	33	48.5	461	2	T39889	protein kinase - f
265	33	48.5	140	2	S41577	lysosome (EC 3.2.1	338	33	48.5	467	2	H69109	heat shock protein
266	33	48.5	150	2	T36673	hypothetical prote	339	33	48.5	468	2	F87359	leucine aminopepti
267	33	48.5	173	1	Y0ECPE	fimbrial protein p	340	33	48.5	492	2	T16659	hypothetical prote
268	33	48.5	173	2	I41201	E-minor fimbrial p	341	33	48.5	495	2	AC3237	monooxygenase Atu6
269	33	48.5	173	2	S25210	papG protein - Esc	342	33	48.5	499	2	C70405	hypothetical prote
270	33	48.5	174	2	JC1294	papG fimbrial prot	343	33	48.5	503	2	AC3296	proline/betaine tr
271	33	48.5	175	2	C27743	papG fimbrial prot	344	33	48.5	505	2	A24707	cytochrome-c oxida
272	33	48.5	196	2	S06613	chorion protein si	345	33	48.5	508	2	E84089	probable transport
273	33	48.5	208	2	B84920	hypothetical prote	346	33	48.5	513	2	S58985	cytochrome-c oxida
274	33	48.5	274	2	T05721	germin-like protei	347	33	48.5	516	2	S52839	cytochrome-c oxida
275	33	48.5	219	1	TEBEG6	tegument protein -	348	33	48.5	519	2	T11129	cytochrome-c oxida
276	33	48.5	219	1	TEBEKA	tegument protein -	349	33	48.5	525	2	S26034	cytochrome-c oxida
277	33	48.5	219	2	A70734	probable cutinase	350	33	48.5	527	1	Q8SEUB	sflB protein - Ssr
278	33	48.5	222	2	T49223	uclacyanin 3 limpo	351	33	48.5	533	2	F75124	abc transporter in
279	33	48.5	229	2	C82117	conserved hypothet	352	33	48.5	536	2	H86132	hypothetical prote
280	33	48.5	236	2	T35095	probable membrane	353	33	48.5	538	2	B81435	probable iron-upta
281	33	48.5	236	2	J00061	hypothetical 25K p	354	33	48.5	554	2	P86927	glucose-6-phosphat
282	33	48.5	246	2	A84921	hypothetical prote	355	33	48.5	556	2	A70560	hypothetical prote
283	33	48.5	251	2	E72578	hypothetical prote	356	33	48.5	562	2	E95366	probable oxidoredu
284	33	48.5	260	2	A38114	resolvase rsd - Sa	357	33	48.5	584	1	S65587	ABC transporter sc
285	33	48.5	279	2	T37010	probable DNA-bindi	358	33	48.5	589	2	A29476	muscarinic acetylch
286	33	48.5	285	2	E87517	pantoate-beta-alan	359	33	48.5	589	2	B29514	muscarinic acetylch
287	33	48.5	286	2	B70632	CDPdiacylglycerol-	360	33	48.5	590	2	S10128	muscarinic acetylch
288	33	48.5	300	2	T36768	secreted serine pr	361	33	48.5	590	2	S01114	muscarinic acetylch
289	33	48.5	300	2	A13101	transcription regu	362	33	48.5	590	2	S47572	muscarinic acetylch
290	33	48.5	308	2	C83586	probable transcript	363	33	48.5	621	2	AB3317	ABC transporter Ar
291	33	48.5	314	2	A87621	hypothetical prote	364	33	48.5	631	2	C90552	lipoprotein (limpo
292	33	48.5	326	2	H70782	probable ompA prot	365	33	48.5	641	2	T07668	starch synthase (E
293	33	48.5	331	2	F75268	hypothetical prote	366	33	48.5	644	2	E88487	protein C13B9.4 li
294	33	48.5	332	2	A96185	probable transcript	367	33	48.5	657	2	A50431	probable exported
295	33	48.5	333	2	E84823	homeotic protein e	368	33	48.5	678	2	S54308	DNA binding protei
296	33	48.5	336	2	D71474	probable muramoyl-	369	33	48.5	683	2	T05149	protein kinase hom
297	33	48.5	336	2	B90574	ABC transporter pe	370	33	48.5	685	2	A83499	probable lipoxigen
298	33	48.5	337	2	S54402	SPX protein - Syn	371	33	48.5	687	2	B70515	hypothetical prote
299	33	48.5	339	2	A83062	hypothetical prote	372	33	48.5	688	2	G82879	translation elonga
300	33	48.5	353	2	E82570	hypothetical prote	373	33	48.5	692	2	S56849	probable membrane
301	33	48.5	357	2	T25499	hypothetical prote	374	33	48.5	695	2	C86731	copper-potassium t
302	33	48.5	362	2	T52038	probable phytochel	375	33	48.5	701	2	T15942	hypothetical prote
303	33	48.5	363	2	H97176	lysophospholipase	376	33	48.5	708	2	T50064	probable transcript
304	33	48.5	368	2	A81128	PTS fructose-speci	377	33	48.5	737	2	H81070	lactoferrin-bindin
305	33	48.5	368	2	AH1488	PTS fructose-speci	378	33	48.5	741	2	D81798	lactoferrin-bindin
306	33	48.5	371	2	B64044	rod shape-determi	379	33	48.5	749	2	E75422	ferrous iron trans
307	33	48.5	374	2	S69699	hypothetical prote	380	33	48.5	753	2	G70184	methyl-accepting c
308	33	48.5	379	2	A49614	CyH protein - Rhi	381	33	48.5	777	2	A35966	chemotaxis protein
309	33	48.5	379	2	A83225	ABC transporter, m	382	33	48.5	790	2	F75372	c-di-GMP phosphodi
310	33	48.5	382	1	SYECLA	lipid A-disacchari	383	33	48.5	837	2	B89583	hypothetical prote
311	33	48.5	382	2	H85502	lipid A-disacchari	384	33	48.5	868	2	AE1953	hypothetical prote
312	33	48.5	382	2	H90651	lipid A-disacchari	385	33	48.5	885	2	C87031	Cell division prot
313	33	48.5	396	2	B75290	hypothetical prote	386	33	48.5	992	2	T27479	hypothetical prote
314	33	48.5	402	2	T04348	endospore specific	387	33	48.5	1088	2	D82246	probable chitinase
315	33	48.5	402	2	C83508	two-component sens	388	33	48.5	1118	2	B83018	conserved hypothet
316	33	48.5	404	2	G69332	heterodisulfide re	389	33	48.5	1120	2	B83498	conserved hypothet
317	33	48.5	407	1	BVBYN1	RNAI protein - yea	390	33	48.5	1195	2	S38174	probable purine nu
318	33	48.5	414	2	B97739	hypothetical prote	391	33	48.5	1235	2	D32433	VSG expression sit
319	33	48.5	415	2	E71677	hypothetical prote	392	33	48.5	1462	1	B36182	protein-tyrosine-p
320	33	48.5	416	2	F83197	probable porin PA3	393	33	48.5	1627	1	IUYMAP	adhesin P1 precurs
321	33	48.5	421	2	T35205	citrate synthase-1	394	33	48.5	1635	2	A41480	adhesin P1, group

395	33	48.5	1696	2	T00057	hypothetical prote	468	32	47.1	217	2	D69271	hypothetical prote
396	33	48.5	1784	2	T02844	cdcl6-related prot	469	32	47.1	218	1	S23664	heib protein - Rho
397	33	48.5	1865	2	G86152	T7123.15 protein -	470	32	47.1	218	2	G87574	rhomboid family pr
398	33	48.5	1955	2	T41170	probable 1,3-beta-	471	32	47.1	231	2	S61235	probable integral
399	33	48.5	1998	2	T08822	nonstructural poly	472	32	47.1	236	2	B83144	conserved hypotet
400	33	48.5	2105	2	T18968	probable serine-ty	473	32	47.1	242	2	T34951	hypothetical prote
401	33	48.5	2655	2	D96595	probable acetyl-Co	474	32	47.1	244	2	B64014	hypothetical prote
402	32.5	47.8	130	2	G70813	probable pqs prot	475	32	47.1	249	2	A37280	C/EBP-related prot
403	32.5	47.8	290	2	F87514	transcription regu	476	32	47.1	249	2	T49993	hypothetical prote
404	32.5	47.8	328	2	A97385	hypothetical prote	477	32	47.1	250	2	T36868	hypothetical prote
405	32.5	47.8	328	2	A12602	components of type	478	32	47.1	251	2	T44070	conserved hypotet
406	32.5	47.8	331	2	D82972	probable lipolytic	479	32	47.1	253	2	AB3586	integral membrane
407	32.5	47.8	368	2	A11291	glycerol dehydroge	480	32	47.1	254	1	PAECGB	phosphatidylglycer
408	32.5	47.8	368	2	AG1663	glycerol dehydroge	481	32	47.1	254	2	C90860	phosphatidylglycer
409	32.5	47.8	393	2	T35894	hypothetical prote	482	32	47.1	254	2	C85759	periplasmic fimbri
410	32.5	47.8	417	2	AE3199	conserved hypotet	483	32	47.1	257	2	AH0859	phosphatidylglycer
411	32.5	47.8	550	2	S75122	acetylactate synth	484	32	47.1	263	2	T12407	NADH2 dehydrogen
412	32.5	47.8	1013	2	B70841	probable helz prot	485	32	47.1	263	2	H69336	cell division inhi
413	32	47.1	36	2	S73127	hypothetical prote	486	32	47.1	265	2	E95926	hypothetical membr
414	32	47.1	70	2	H69840	conserved hypotet	487	32	47.1	280	2	H95352	hypothetical prote
415	32	47.1	71	2	S45608	light-harvesting p	488	32	47.1	280	2	H70942	hypothetical prote
416	32	47.1	73	2	AB2468	secretory protein	489	32	47.1	282	2	F71077	hypothetical prote
417	32	47.1	89	2	A10221	flagellar biosynth	490	32	47.1	282	2	AG2954	hypothetical prote
418	32	47.1	93	2	T10379	hypothetical prote	491	32	47.1	285	2	F98328	sugar ABC transpor
419	32	47.1	95	2	S72597	hypothetical prote	492	32	47.1	285	2	AE3024	hypothetical prote
420	32	47.1	99	1	R6MXL2	ribosomal protein	493	32	47.1	286	2	AH1062	probable phage cap
421	32	47.1	102	2	D90915	hypothetical prote	494	32	47.1	286	2	E9516	hypothetical prote
422	32	47.1	102	2	A85764	hypothetical prote	495	32	47.1	290	2	AE0677	probable regulator
423	32	47.1	104	2	G86639	hypothetical prote	496	32	47.1	294	2	S36932	chitinase (EC 3.2.
424	32	47.1	106	2	T06479	proline/leucine-ri	497	32	47.1	295	2	E70022	transcription regu
425	32	47.1	107	2	T35523	probable small sec	498	32	47.1	297	2	A69353	branched-chain ami
426	32	47.1	109	2	A54678	monocyte chemotact	499	32	47.1	299	1	RPECGP	glycerol-3-phospha
427	32	47.1	109	2	B70037	hypothetical prote	500	32	47.1	300	2	H86636	hypothetical prote
428	32	47.1	111	2	T05924	probable cysteine	501	32	47.1	303	2	AG3303	putrescine transpo
429	32	47.1	111	2	D65048	hypothetical prote	502	32	47.1	305	2	B96840	hypothetical prote
430	32	47.1	111	2	A91072	hypothetical prote	503	32	47.1	305	2	D83699	hypothetical prote
431	32	47.1	115	2	JC6511	motilin precursor	504	32	47.1	307	2	B82394	hypothetical prote
432	32	47.1	115	2	D64914	hypothetical prote	505	32	47.1	310	2	AE2775	conserved hypotet
433	32	47.1	121	2	S34233	hypothetical prote	506	32	47.1	310	2	C97555	xylanase [imported
434	32	47.1	134	2	S71274	blue copper-bindin	507	32	47.1	317	2	D98260	inner membrane pro
435	32	47.1	136	2	B45893	T-cell receptor al	508	32	47.1	322	2	S75428	UDP-N-acetylglucos
436	32	47.1	137	2	T25123	hypothetical prote	509	32	47.1	322	2	E84330	hypothetical prote
437	32	47.1	138	2	A29479	thyrotropin beta c	510	32	47.1	326	2	E82764	hypothetical prote
438	32	47.1	139	2	C83526	hypothetical prote	511	32	47.1	329	1	BVECEX	pyridoxal phosphat
439	32	47.1	141	2	T08646	blue copper-bindin	512	32	47.1	330	2	B82822	NADH2 dehydrogen
440	32	47.1	148	2	S39403	hupG protein - Bra	513	32	47.1	336	2	H72618	hypothetical prote
441	32	47.1	149	2	E85067	hypothetical prote	514	32	47.1	346	2	B75439	hypothetical prote
442	32	47.1	151	2	T36081	probable integral	515	32	47.1	348	2	T36890	probable iron-side
443	32	47.1	152	2	F72702	hypothetical prote	516	32	47.1	349	2	D98170	sugar ABC transpor
444	32	47.1	154	2	F72741	hypothetical prote	517	32	47.1	351	2	T25448	hypothetical prote
445	32	47.1	157	2	AC2320	hypothetical prote	518	32	47.1	354	2	S77275	glyceraldehyde-3-p
446	32	47.1	159	2	S72846	hypothetical prote	519	32	47.1	354	2	C82110	probable permease
447	32	47.1	162	2	G84363	hypothetical prote	520	32	47.1	355	2	B70678	hypothetical prote
448	32	47.1	166	2	C91031	probable fimbrial	521	32	47.1	370	2	I51199	activin beta B sub
449	32	47.1	169	2	A83134	conserved hypotet	522	32	47.1	370	2	G95976	succinyltransferas
450	32	47.1	172	2	S71273	blue copper-bindin	523	32	47.1	371	2	AH0147	outer membrane pro
451	32	47.1	175	2	AH0096	conserved hypotet	524	32	47.1	372	2	S34803	exoh protein - Rhi
452	32	47.1	179	2	B83484	probable signal pe	525	32	47.1	372	2	AG2049	permease protein o
453	32	47.1	184	2	T33006	hypothetical prote	526	32	47.1	375	2	I38879	corticotropin rele
454	32	47.1	190	2	S66981	ribosomal protein	527	32	47.1	379	2	G70918	hypothetical prote
455	32	47.1	192	2	AB0726	conserved hypotet	528	32	47.1	385	2	AH0083	probable iron-cont
456	32	47.1	194	2	I50710	fibroblast growth	529	32	47.1	386	2	T08353	hypothetical prote
457	32	47.1	197	2	D85875	probable minor fil	530	32	47.1	388	2	C95175	membrane protein l
458	32	47.1	204	2	D87695	hypothetical prote	531	32	47.1	388	2	D98041	conserved hypotet
459	32	47.1	212	2	A84359	hypothetical prote	532	32	47.1	394	2	T36852	probable integral
460	32	47.1	213	2	G82723	partition protein	533	32	47.1	398	2	T40074	hypothetical zf-C3
461	32	47.1	214	2	AD3279	hypothetical prote	534	32	47.1	411	2	A55610	corticotropin-rele
462	32	47.1	214	2	T44107	conserved hypotet	535	32	47.1	415	2	S39535	corticotropin-rele
463	32	47.1	214	2	G89763	conserved hypotet	536	32	47.1	415	2	I58144	corticotropin-rele
464	32	47.1	214	2	B86156	hypothetical prote	537	32	47.1	419	2	S23425	Preproenke translo
465	32	47.1	215	2	T06909	glutamine amidotra	538	32	47.1	419	2	T51415	Carboxylesterase-1
466	32	47.1	217	2	H69669	choline ABC transp	539	32	47.1	423	1	TFV2E	transforming prote
467	32	47.1	217	2	AI0987	probable lipoprote	540	32	47.1	423	2	H83603	hypothetical prote

541	32	47.1	424	2	H84134	multidrug resistanc	614	32	47.1	618	2	D87651	prolyl oligopeptid
542	32	47.1	425	2	AC2092	hypothetical prote	615	32	47.1	619	2	G71861	primosomal protein
543	32	47.1	429	1	B41902	arsenical pump mem	616	32	47.1	619	2	C64568	primosomal protein
544	32	47.1	429	1	C41903	arsenical pump mem	617	32	47.1	628	2	C82842	bioethnetic argin
545	32	47.1	429	2	A01894	serine proteinase	618	32	47.1	628	2	F70556	hypothetical prote
546	32	47.1	430	2	G89962	hypothetical prote	619	32	47.1	635	2	S75668	DNA primase IEC 2.
547	32	47.1	430	2	A56726	corticolliberin rec	620	32	47.1	652	2	B29900	fasciclin I precu
548	32	47.1	430	2	E82502	conserved hypothet	621	32	47.1	653	2	T34976	probable integrat
549	32	47.1	431	2	A01237	arsenic efflux pum	622	32	47.1	654	2	T34960	cell division prot
550	32	47.1	431	2	A11599	arsenic efflux pum	623	32	47.1	662	2	T41512	us snrnp-like RNA
551	32	47.1	431	2	I49149	CRP receptor - nou	624	32	47.1	678	2	G65104	hypothetical 72.8
552	32	47.1	431	2	I49279	sauvagine/corticot	625	32	47.1	678	2	D91132	probable glycosyla
553	32	47.1	434	2	G97610	hypothetical metab	626	32	47.1	678	2	G85977	probable glycosyla
554	32	47.1	434	2	AC2833	hypothetical prote	627	32	47.1	683	2	AD0899	probable exported
555	32	47.1	435	2	G69779	arsenical pump mem	628	32	47.1	685	2	AC0537	ferichrome transp
556	32	47.1	435	2	G84024	arsenic efflux pum	629	32	47.1	685	2	T18964	hypothetical prote
557	32	47.1	436	2	E95149	uracil permease li	630	32	47.1	688	2	T21641	hypothetical prote
558	32	47.1	439	2	D98017	uracil permease li	631	32	47.1	697	2	T49119	hypothetical prote
559	32	47.1	440	2	JC2532	secretin receptor	632	32	47.1	704	2	G81051	hypothetical prote
560	32	47.1	441	2	C91250	hypothetical prote	633	32	47.1	712	2	AG1888	Na+/H+ antiporter
561	32	47.1	441	2	C70780	probable pPG prote	634	32	47.1	720	2	T20016	hypothetical prote
562	32	47.1	443	2	A48260	corticolliberin rec	635	32	47.1	720	2	T15231	gemline RNA helic
563	32	47.1	446	2	A83213	probable ATP-depen	636	32	47.1	740	2	E81823	hypothetical prote
564	32	47.1	450	2	A10909	probable membrane	637	32	47.1	741	2	A83271	hypothetical prote
565	32	47.1	452	2	E83482	probable MFS trans	638	32	47.1	749	2	AD0345	probable phosphate
566	32	47.1	452	2	T35729	hypothetical prote	639	32	47.1	783	2	A48998	nucleolar protein
567	32	47.1	452	2	G95347	Protein imported	640	32	47.1	806	2	T35640	probable sensor ki
568	32	47.1	455	2	A70717	hypothetical prote	641	32	47.1	814	2	F75464	probable NADH-ubiq
569	32	47.1	456	2	E83750	gluconate permease	642	32	47.1	829	2	A99960	hypothetical prote
570	32	47.1	460	2	S39715	conserved hypothet	643	32	47.1	840	2	A12153	hypothetical prote
571	32	47.1	463	2	D87624	sodium-galactoside	644	32	47.1	855	2	A48168	proliferating-cell
572	32	47.1	465	2	T34984	probable integral	645	32	47.1	899	2	G71453	hypothetical prote
573	32	47.1	469	2	T11910	NADH2 dehydrogenas	646	32	47.1	905	2	S75035	H+-exporting ATPas
574	32	47.1	470	2	A87547	vanillin dehydroge	647	32	47.1	947	2	B90934	probable molybdate
575	32	47.1	472	2	S77364	DNA damage-inducib	648	32	47.1	963	2	E85839	probable regulator
576	32	47.1	475	2	A95376	probable Arcdi arg	649	32	47.1	974	2	E59434	Rho GTPase activat
577	32	47.1	482	2	S08384	modulation protein	650	32	47.1	1029	2	T41386	RANBP7/importin-be
578	32	47.1	482	2	S10133	modulation protein	651	32	47.1	1062	2	S46367	protein kinase CDC
579	32	47.1	490	2	C70146	glutamate-tRNA lig	652	32	47.1	1160	1	DJEC3A	DNA-directed DNA p
580	32	47.1	497	2	AG1768	hypothetical prote	653	32	47.1	1160	2	B85503	DNA polymerase III
581	32	47.1	498	2	S82626	protein disulfide-	654	32	47.1	1160	2	B90652	DNA polymerase III
582	32	47.1	499	2	C90317	hypothetical prote	655	32	47.1	1208	2	C82779	hemolysin-type cal
583	32	47.1	509	2	T11043	cytochrome-c oxida	656	32	47.1	1220	2	T18291	patched protein -
584	32	47.1	510	2	T13882	cytochrome-c oxida	657	32	47.1	1324	2	T17468	peptide-synthetase
585	32	47.1	510	2	S76707	hypothetical prote	658	32	47.1	1366	2	B84924	hypothetical prote
586	32	47.1	511	2	E84685	probable seed stor	659	32	47.1	1388	2	T34157	hypothetical prote
587	32	47.1	512	2	T11154	cytochrome-c oxida	660	32	47.1	1434	2	T30172	transmembrane prot
588	32	47.1	512	2	S60622	cytochrome-c oxida	661	32	47.1	1442	2	T18538	patched protein -
589	32	47.1	513	2	I39534	hypothetical prote	662	32	47.1	1469	2	T50210	probable ABC trans
590	32	47.1	519	2	C87350	major facilitator	663	32	47.1	1844	1	RRWPTM	genome polyprotein
591	32	47.1	519	2	B87353	hypothetical prote	664	32	47.1	1844	2	S01956	hypothetical prote
592	32	47.1	527	2	G69635	PTS arbutin-like e	665	32	47.1	1887	2	S61703	fatty-acid synthas
593	32	47.1	528	2	AG1937	glucose-6-phosphat	666	32	47.1	1985	2	S19151	hypothetical prote
594	32	47.1	531	2	T14640	cytochrome P450 CY	667	32	47.1	3172	2	S22012	erythronolide synt
595	32	47.1	534	2	F72679	hypothetical prote	668	32	47.1	3178	2	S13595	6-deoxyerythronoli
596	32	47.1	543	2	S62012	probable membrane	669	31.5	46.3	100	2	C64575	hypothetical prote
597	32	47.1	552	2	T27424	hypothetical prote	670	31.5	46.3	205	2	AT2614	RhCB family transp
598	32	47.1	558	2	T33649	hypothetical prote	671	31.5	46.3	212	2	AD1958	hypothetical prote
599	32	47.1	566	2	T03855	hypothetical prote	672	31.5	46.3	224	2	H97396	dihydrodipicolinat
600	32	47.1	576	2	H70961	hypothetical prote	673	31.5	46.3	226	2	T30478	hypothetical prote
601	32	47.1	576	2	T36729	probable serine/th	674	31.5	46.3	318	2	D87506	glycosyl transfera
602	32	47.1	583	2	B82744	conserved hypothet	675	31.5	46.3	376	2	E70680	probable proB - My
603	32	47.1	589	2	E93736	sulfite reductase	676	31.5	46.3	571	2	H87600	hypothetical prote
604	32	47.1	589	2	AD2594	hypothetical prote	677	31.5	46.3	680	2	H87091	probable NAD synth
605	32	47.1	595	2	T41007	atp dependent heli	678	31	45.6	51	2	S17514	70K protein - rabb
606	32	47.1	599	2	JH0259	prostaglandin-endo	679	31	45.6	22	2	G82455	hypothetical prote
607	32	47.1	599	2	T24333	hypothetical prote	680	31	45.6	77	2	S47941	VMA21 protein - ye
608	32	47.1	600	2	C83142	hypothetical prote	681	31	45.6	80	2	S34219	acclimation protei
609	32	47.1	604	2	T15132	ATP-dependent RNA	682	31	45.6	84	1	WMBPBH	gene h protein - p
610	32	47.1	604	2	C87818	protein glh-1 limp	683	31	45.6	84	2	T04475	probable acclimati
611	32	47.1	604	2	T04351	viviparous-14 prot	684	31	45.6	90	2	PC2137	hypothetical 90 pr
612	32	47.1	613	2	C83496	glutathione-regula	685	31	45.6	90	2	PC2136	Lim1 protein - tru
613	32	47.1	613	2	JC7762	SOX-3 protein - gu	686	31	45.6	95	2	D84327	hypothetical prote



687 31 45.6 99 2 B81051  
688 31 45.6 99 2 D81825  
689 31 45.6 101 2 S39490  
690 31 45.6 104 2 E75228  
691 31 45.6 104 2 T30750  
692 31 45.6 110 2 A64093  
693 31 45.6 111 2 D85916  
694 31 45.6 112 2 JQ2290  
695 31 45.6 115 1 A33323  
696 31 45.6 116 2 T02044  
697 31 45.6 116 2 S37659  
698 31 45.6 117 2 D72650  
699 31 45.6 118 2 S59330  
700 31 45.6 119 2 A72687  
701 31 45.6 120 2 E07038  
702 31 45.6 121 2 JQ2291  
703 31 45.6 121 2 B84375  
704 31 45.6 124 2 AB1977  
705 31 45.6 125 2 AC1862  
706 31 45.6 128 2 S12900  
707 31 45.6 133 2 H84290  
708 31 45.6 133 2 T35848  
709 31 45.6 143 2 S42579  
710 31 45.6 144 2 B95905  
711 31 45.6 146 2 S11398  
712 31 45.6 149 2 B81002  
713 31 45.6 149 2 B82017  
714 31 45.6 152 2 H71129  
715 31 45.6 154 2 T50898  
716 31 45.6 156 2 B44048  
717 31 45.6 157 2 AD2320  
718 31 45.6 158 2 H87711  
719 31 45.6 159 2 T30745  
720 31 45.6 160 2 T34665  
721 31 45.6 161 2 B84937  
722 31 45.6 161 2 D71892  
723 31 45.6 163 2 S06152  
724 31 45.6 163 2 F81339  
725 31 45.6 163 2 AG2365  
726 31 45.6 169 2 AF0148  
727 31 45.6 169 2 AI2391  
728 31 45.6 173 2 C41281  
729 31 45.6 173 2 E41281  
730 31 45.6 173 2 A41281  
731 31 45.6 173 2 B41281  
732 31 45.6 173 2 F41281  
733 31 45.6 173 2 H41281  
734 31 45.6 176 2 C55208  
735 31 45.6 178 2 S68681  
736 31 45.6 181 2 F75561  
737 31 45.6 182 2 E64678  
738 31 45.6 182 2 C71839  
739 31 45.6 190 2 S74507  
740 31 45.6 192 2 G75419  
741 31 45.6 193 2 S24295  
742 31 45.6 194 2 A69760  
743 31 45.6 195 2 B82443  
744 31 45.6 196 2 E75612  
745 31 45.6 197 2 C70601  
746 31 45.6 199 2 G69935  
747 31 45.6 199 2 AB0216  
748 31 45.6 201 2 AI0409  
749 31 45.6 201 2 D95880  
750 31 45.6 204 2 S62203

ALIGNMENTS

RESULT 1  
AF3275  
transporter, dme family BMEI0187 [imported] - Brucella melitensis (strain 16M)  
C:Species: Brucella melitensis

C:Date: 01-Feb-2002 #sequence\_revision 01-Feb-2002 #text\_change 09-Jul-2004  
C:Accession: AF3275  
R:DelVecchio, V.G.; Kapatral, V.; Redkar, R.J.; Patra, G.; Mijer, C.; Los, T.; Ivanova,  
.; Mazur, M.; Goltzman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letes  
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002  
A:Title: The genome sequence of the facultative intracellular pathogen Brucella meliten  
A:Reference number: AD3252; PMID:11756688  
A:Accession: AF3275  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-297 <KUR>  
A:Cross-references: UNIPROT:O8YJ99; GB:AE008917; PIDN:AAL51369.1; PID:gl7982069; GSPDB:  
A:Experimental source: strain 16M  
C:Genetics:  
A:Gene: BMEI0187  
A:Map position: 1  
  
Query Match 61.8%; Score 42; DB 2; Length 297;  
Best Local Similarity 81.8%; Pred. No. 16;  
Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
  
QY 3 AAALLPVL 13  
Db 160 AAALLPVL 170  
  
RESULT 2  
T05209  
hypothetical protein F24J7.50 - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 09-Jul-2004  
A:Accession: T05209  
R:Bevan, M.; Vitale, D.; Lignori, R.; Argirou, A.; De Simone, V.; Bancroft, I.; Mewes,  
submitted to the Protein Sequence Database, February 1998  
A:Reference number: Z15403  
A:Accession: T05209  
A:Molecule type: DNA  
A:Residues: 1-939 <BEV>  
A:Cross-references: UNIPROT:O49467; EMBL:AL021768; GSPDB:GN000062; ATSP:F24J7.50  
A:Experimental source: cultivar Columbia; BAC clone F24J7  
C:Genetics:  
A:Gene: ATSP:F24J7.50  
A:Map position: 4  
A:Introns: 278/3; 299/3; 326/2; 336/2; 376/3; 401/3; 453/1; 492/3; 578/2; 635/3; 663/3;  
  
Query Match 61.8%; Score 42; DB 2; Length 939;  
Best Local Similarity 69.2%; Pred. No. 45;  
Matches 9; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
  
QY 1 KKAALLPVL 13  
Db 424 KKAALLPVL 436  
  
RESULT 3  
D72482  
probable antibiotic transport-associated protein APE2500 - Aeropyrum pernix (strain K1)  
C:Species: Aeropyrum pernix  
C:Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 09-Jul-2004  
A:Accession: D72482  
R:Kawarabayasi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takai  
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; i  
DNA Res. 6, 83-101, 1999  
A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr  
A:Reference number: A72450; MUID:99310339; PMID:10382966  
A:Accession: D72482  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-1172 <KAW>  
A:Cross-references: UNIPROT:Q9Y8Y4; DDBJ:AF000064; NID:gs105945; PIDN:BAA81516.1; PID:d  
A:Experimental source: strain K1  
C:Genetics:  
A:Gene: APE2500

Query Match 60.3%; Score 41; DB 2; Length 1172;  
Best Local Similarity 69.2%; Pred. No. 81;  
Matches 9; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 KAAVLLPVLLAA 14  
||| ||||| : :  
DB 848 KAAVLLPVLLVVA 860

RESULT 4  
T09155  
lipid transfer protein - spinach  
C:Species: Spinacia oleracea (spinach)  
C:Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 09-Jul-2004  
C:Accession: T09155  
R:Bernhard, W.R.; Thoma, S.; Borella, J.; Somerville, C.R.  
Plant Physiol. 95, 164-170, 1991  
A:Title: Isolation of a cDNA clone for spinach lipid transfer protein and evidence that  
A:Reference number: 216594  
A:Accession: T09155  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-117 <BER>  
A:Cross-references: UNIPROT:P10976; EMBL:M58635; NID:gi70116; PID:gi70117  
C:Superfamily: phospholipid transfer protein

Query Match 58.8%; Score 40; DB 2; Length 117;  
Best Local Similarity 60.0%; Pred. No. 14;  
Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 KKAARVLLPVLLAA 15  
||| ||||| : : : :  
DB 8 KLACAVLLCIVVAP 22

RESULT 5  
AI2821  
conserved hypothetical protein Atul998 [imported] - Agrobacterium tumefaciens (strain C58)  
C:Species: Agrobacterium tumefaciens  
C:Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 09-Jul-2004  
C:Accession: AI2821  
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L.;  
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClell  
; Karp, P.; Romero, P.; Zhang, S.  
Science 294, 2317-2323, 2001  
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,  
ster, E.W.  
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.  
A:Reference number: AB2577; MUID:21608550; PMID:11743193  
A:Accession: AI2821  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-158 <KUR>  
A:Cross-references: UNIPROT:Q8UDX0; GB:AE008688; PIDN:AAL42991.1; PID:gi7740452; GSPDB:G  
A:Experimental source: strain C58 (Dupont)  
C:Genetics:  
A:Gene: Atul998  
A:Map position: circular chromosome

Query Match 58.8%; Score 40; DB 2; Length 159;  
Best Local Similarity 69.2%; Pred. No. 19;  
Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 AAAVLLPVLLAA 15  
||| ||||| : : : :  
DB 11 AALALAPALLAA 23

RESULT 6  
A97600  
hypothetical protein AGR\_C\_3631 [imported] - Agrobacterium tumefaciens (strain C58, Cere  
C:Species: Agrobacterium tumefaciens

Query Match 60.3%; Score 41; DB 2; Length 1172;  
Best Local Similarity 69.2%; Pred. No. 81;  
Matches 9; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 KAAVLLPVLLAA 14  
||| ||||| : :  
DB 848 KAAVLLPVLLVVA 860

RESULT 4  
T09155  
lipid transfer protein - spinach  
C:Species: Spinacia oleracea (spinach)  
C:Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 09-Jul-2004  
C:Accession: T09155  
R:Bernhard, W.R.; Thoma, S.; Borella, J.; Somerville, C.R.  
Plant Physiol. 95, 164-170, 1991  
A:Title: Isolation of a cDNA clone for spinach lipid transfer protein and evidence that  
A:Reference number: 216594  
A:Accession: T09155  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-117 <BER>  
A:Cross-references: UNIPROT:P10976; EMBL:M58635; NID:gi70116; PID:gi70117  
C:Superfamily: phospholipid transfer protein

Query Match 58.8%; Score 40; DB 2; Length 117;  
Best Local Similarity 60.0%; Pred. No. 14;  
Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 KKAARVLLPVLLAA 15  
||| ||||| : : : :  
DB 8 KLACAVLLCIVVAP 22

RESULT 5  
AI2821  
conserved hypothetical protein Atul998 [imported] - Agrobacterium tumefaciens (strain C58)  
C:Species: Agrobacterium tumefaciens  
C:Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 09-Jul-2004  
C:Accession: AI2821  
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L.;  
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClell  
; Karp, P.; Romero, P.; Zhang, S.  
Science 294, 2317-2323, 2001  
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,  
ster, E.W.  
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.  
A:Reference number: AB2577; MUID:21608550; PMID:11743193  
A:Accession: AI2821  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-158 <KUR>  
A:Cross-references: UNIPROT:Q8UDX0; GB:AE008688; PIDN:AAL42991.1; PID:gi7740452; GSPDB:G  
A:Experimental source: strain C58 (Dupont)  
C:Genetics:  
A:Gene: Atul998  
A:Map position: circular chromosome

Query Match 58.8%; Score 40; DB 2; Length 159;  
Best Local Similarity 69.2%; Pred. No. 19;  
Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 AAAVLLPVLLAA 15  
||| ||||| : : : :  
DB 11 AALALAPALLAA 23

RESULT 6  
A97600  
hypothetical protein AGR\_C\_3631 [imported] - Agrobacterium tumefaciens (strain C58, Cere  
C:Species: Agrobacterium tumefaciens

C:Date: 30-Sep-2001 #sequence\_revision 30-Sep-2001 #text\_change 09-Jul-2004  
C:Accession: A97600  
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,  
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;  
Science 294, 2323-2328, 2001  
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tur  
A:Reference number: A97359; MUID:21608551; PMID:11743194  
A:Accession: A97600  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-163 <KUR>  
A:Cross-references: UNIPROT:Q8UDX0; GB:AE007869; PIDN:AAK87754.1; PID:gi51517122; GSPDB:G  
C:Genetics:  
A:Gene: AGR\_C\_3631  
A:Map position: circular chromosome

Query Match 58.8%; Score 40; DB 2; Length 163;  
Best Local Similarity 69.2%; Pred. No. 19;  
Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 AAAVLLPVLLAA 15  
||| ||||| : : : :  
DB 16 AALALAPALLAA 28

RESULT 7  
A72698  
hypothetical protein APE1001 - Aeropyrum pernix (strain K1)  
C:Species: Aeropyrum pernix  
C:Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 09-Jul-2004  
C:Accession: A72698  
R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takat  
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K  
DNA Res. 6, 83-101, 1999  
A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropy  
A:Reference number: A72450; MUID:99310339; PMID:10382966  
A:Accession: A72698  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-390 <KAW>  
A:Cross-references: UNIPROT:Q9YDB2; DBJ:AP000060; NID:gs104188; PIDN:BAA79985.1; PID:gi  
A:Experimental source: strain K1  
C:Genetics:  
A:Gene: APE1001

Query Match 58.8%; Score 40; DB 2; Length 390;  
Best Local Similarity 75.0%; Pred. No. 43;  
Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 AAVALPVLLAA 15  
||| ||||| : : : :  
DB 180 AAKLLPVLLAA 191

RESULT 8  
AF2635  
exopolysaccharide production protein [imported] - Agrobacterium tumefaciens (strain C58,  
C:Species: Agrobacterium tumefaciens  
C:Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 09-Jul-2004  
C:Accession: AF2635  
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L  
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClell  
; Karp, P.; Romero, P.; Zhang, S.  
Science 294, 2317-2323, 2001  
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,  
ster, E.W.  
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.  
A:Reference number: AB2577; MUID:21608550; PMID:11743193  
A:Accession: AF2635  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-431 <KUR>  
A:Cross-references: UNIPROT:Q8UI19; GB:AE008688; PIDN:AAL41500.1; PID:gi7738828; GSPDB:G

A:Experimental source: strain C58 (Dupont)

C:Genetics:

A:Gene: exoQ

A:Map position: circular chromosome

Query Match 58.8%; Score 40; DB 2; Length 431;

Best Local Similarity 75.0%; Pred. No. 47;

Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 KAAAVLLPVLLA 13

|||||

227 KTAAMLLPVTLA 238

RESULT 9

E97417

exoQ-like protein (AF039306) [imported] - Agrobacterium tumefaciens (strain C58, Cereon)

C:Species: Agrobacterium tumefaciens

C:Date: 30-Sep-2001 #sequence\_revision 30-Sep-2001 #text\_change 09-Jul-2004

C:Accession: E97417

R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorollo, B.; Goldman,

A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;

Science 294, 2323-2328, 2001

A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum

A:Reference number: A97359; MUID:21608551; PMID:11743194

A:Accession: E97417

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-435 <KUR>

A:Cross-references: UNIPROT:Q8U119; GB:AE007869; PIDN:AAK86294.1; PID:g15155408; GSPDB:C

C:Genetics:

A:Gene: AGR\_C\_854

A:Map position: circular chromosome

Query Match

Best Local Similarity 58.8%; Score 40; DB 2; Length 435;

Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 KAAAVLLPVLLA 13

|||||

231 KTAAMLLPVTLA 242

RESULT 10

I48673

matrix metalloproteinase (EC 3.4.24.-) membrane type precursor - mouse

C:Species: Mus musculus (house mouse)

C:Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 22-Jun-1999

C:Accession: I48673

R:Okada, A.; Bellocq, J.P.; Rouyer, N.; Chenard, M.P.; Rio, M.C.; Chambon, P.; Basset, P.

Proc. Natl. Acad. Sci. U.S.A. 92, 2730-2734, 1995

A:Title: Membrane-type matrix metalloproteinase (MT-MMP) gene is expressed in stromal ce

A:Reference number: I38046; MUID:95224014; PMID:7708715

A:Accession: I48673

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-582 <RES>

A:Cross-references: EMBL:X83536; NID:g804999; PIDN:CAA58520.1; PID:g805000

C:Superfamily: interstitial collagenase; hemopexin repeat homology; matrix metalloprotei

C:Keywords: hydrolase; metalloproteinase; zinc; zymogen

F:1-23/Domain: signal sequence #status predicted <SIG>

F:24-97/Domain: activation peptide #status predicted <PRO>

F:61-284/Domain: matrix metalloproteinase homology <MMP>

F:313-508/Domain: hemopexin repeat homology <HPN>

F:93-239,243/Binding site: zinc, catalytic (Cys, His, His) (inhibited) #status

F:239,243,249/Binding site: zinc, catalytic (His) (active) #status predicted

F:240/Active site: Glu #status predicted

Query Match

Best Local Similarity 58.8%; Score 40; DB 2; Length 582;

Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 AAAVLLPVLL 12

Db 539 AAAVLLPVLL 548

|||||

RESULT 11

I38028

matrix metalloproteinase 14 (EC 3.4.24.-) membrane type precursor - human

N:Alternate names: matrix metalloproteinase MMP14; membrane type matrix metalloproteina

C:Species: Homo sapiens (man)

C:Date: 17-May-1996 #sequence\_revision 17-May-1996 #text\_change 09-Jul-2004

C:Accession: I38028; G02274; I38046; S78011; S45341; S71384

R:Will, H.; Hinzmann, B.

Eur. J. Biochem. 231, 602-608, 1995

A:Title: CDNA sequence and mRNA tissue distribution of a novel human matrix metalloprot

A:Reference number: I38028; MUID:95377289; PMID:7649159

A:Accession: I38028

A:Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: mRNA

A:Residues: 1-582 <WILL>

A:Cross-references: UNIPROT:P50281; EMBL:Z48481; NID:g963053; PIDN:CAA88372.1; PID:g9630

R:Juo, G.

submitted to the EMBL Data Library, November 1995

A:Reference number: H00963

A:Accession: G02274

A:Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: mRNA

A:Residues: 1-582 <JJO>

A:Cross-references: EMBL:U41078; NID:g1127836; PIDN:AAA83770.1; PID:g1127837

R:Okada, A.; Bellocq, J.P.; Rouyer, N.; Chenard, M.P.; Rio, M.C.; Chambon, P.; Basset, P.

Proc. Natl. Acad. Sci. U.S.A. 92, 2730-2734, 1995

A:Title: Membrane-type matrix metalloproteinase (MT-MMP) gene is expressed in stromal ce

A:Reference number: I38046; MUID:95224014; PMID:7708715

A:Accession: I38046

A:Status: translated from GB/EMBL/DBDJ

A:Molecule type: mRNA

A:Residues: 1-7,'S'

A:Cross-references: EMBL:X83535; NID:g804993; PIDN:CAA58519.1; PID:g804994

R:Seiki, M.

submitted to the EMBL Data Library, January 1994

A:Reference number: S78011

A:Accession: S78011

A:Molecule type: mRNA

A:Residues: 1-7,'S'

A:Cross-references: EMBL:D26512; NID:g793762; PIDN:BAA05519.1; PID:g793763

R:Sato, H.; Takino, T.; Okada, Y.; Cao, J.; Shingawa, A.; Yamamoto, E.; Seiki, M.

Nature 370, 61-65, 1994

A:Title: A matrix metalloproteinase expressed on the surface of invasive tumour cells.

A:Reference number: S45341; MUID:94286011; PMID:8015608

A:Accession: S45341

A:Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-7,'S'

A:Cross-references: EMBL:D26512

R:Sato, H.; Kinoshita, T.; Takino, T.; Nakayama, K.; Seiki, M.

FEBS Lett. 393, 101-104, 1996

A:Title: Activation of a recombinant membrane type 1-matrix metalloproteinase (MT1-MMP)

A:Reference number: S71384; MUID:96397540; PMID:8804434

A:Accession: S71384

A:Molecule type: protein

A:Residues: 112-116 <SAW>

C:Genetics:

A:Gene: GDB:MMP14; MT1-MMP

A:Cross-references: GDB:375731; OMIM:600754

A:Map position: 14q11-14q12

C:Superfamily: interstitial collagenase; hemopexin repeat homology; matrix metalloprote

C:Keywords: glycoprotein; hydrolase; metalloproteinase; zinc; zymogen

F:1-23/Domain: signal sequence #status predicted <SIG>

F:24-97/Domain: activation peptide #status predicted <PRO>

F:61-284/Domain: matrix metalloproteinase homology <MMP>

F:98-582/Product: hinge #status predicted <HNG>

F:285-313/Domain: hinge #status predicted <HNG>

F:314-508/Domain: hemopexin repeat homology <HPN>

F:539-562/Domain: transmembrane #status predicted <TMW>

F;93-239,243,249/Binding site: zinc, catalytic (Cys, His, His) (inhibited) #status  
F;130/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F;239,243,249/Binding site: zinc, catalytic (His) (active) #status predicted  
F;240/Active site: Glu #status predicted  
F;319-508/Disulfide bonds: #status predicted

Query Match 58.8%; Score 40; DB 2; Length 582;  
Best Local Similarity 90.0%; Pred. No. 62;  
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 AAVALLPVLL 12  
||||:|||||  
Db 539 AAVALPVL 548

RESULT 12  
184471  
matrix metalloproteinase (EC 3.4.24.-) membrane type precursor - rat  
N:Alternate names: membrane-type metalloproteinase  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 09-Jul-2004  
C:Accession: I61946  
R:Okada, A.; Bellocq, J.P.; Rouyer, N.; Chenard, M.P.; Rio, M.C.; Chambon, P.; Basset, P.  
Proc. Natl. Acad. Sci. U.S.A. 92, 2730-2734, 1995  
A:Title: Membrane-type matrix metalloproteinase (MT-MMP) gene is expressed in stromal ce  
A:Reference number: 138046; MUID:95224014; PMID:7708715  
A:Accession: I61946  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-582 <RES>  
A:Cross-references: UNIPROT:Q10739; EMBL:X83537; NID:G805012; PIDN:CAA58521.1; PID:G8050  
A:Accession: I61946  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-67, 'M', 69-254, 'A', 256-582 <RE2>  
A:Cross-references: EMBL:X91785; NID:G1001926; PIDN:CAA62897.1; PID:G1001927  
C:Genetics:  
A:Gene: mt-mmp  
C:Superfamily: interstitial collagenase; hemopexin repeat homology; matrix metalloprotei  
C:Keywords: hydrolase; metalloproteinase; zinc; zymogen  
F;1-23/Domain: signal sequence #status predicted <SIG>  
F;24-97/Domain: activation peptide #status predicted <PRO>  
F;61-284/Domain: matrix metalloproteinase homology <MPP>  
F;313-508/Domain: hemopexin repeat homology <PXN>  
F;93-239,243,249/Binding site: zinc, catalytic (Cys, His, His) (inhibited) #status  
F;239,243,249/Binding site: zinc, catalytic (His) (active) #status predicted  
F;240/Active site: Glu #status predicted

Query Match 58.8%; Score 40; DB 2; Length 582;  
Best Local Similarity 90.0%; Pred. No. 62;  
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 AAVALLPVLL 12  
||||:|||||  
Db 539 AAVALPVL 548

RESULT 13  
E90536  
hypothetical protein MYPU\_1970 [imported] - Mycoplasma pulmonis (strain UAB CTIP)  
C:Species: Mycoplasma pulmonis  
C:Date: 24-May-2001 #sequence\_revision 24-May-2001 #text\_change 09-Jul-2004  
C:Accession: E90536  
R:Chambaud, I.; Heilig, R.; Ferris, S.; Barbe, V.; Samson, D.; Galisson, F.; Moszer, I.;  
Nucleic Acids Res. 29, 2145-2153, 2001  
A:Title: The complete genome sequence of the murine respiratory pathogen Mycoplasma pulm  
A:Reference number: A99512; MUID:21267165; PMID:11353084  
A:Accession: E90536  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-606 <KUP>  
A:Cross-references: UNIPROT:Q98R13; GB:AL445566; PID:G14089610; PIDN:CAC13370.1; GSPDB:G  
A:Experimental source: strain UAB CTIP

C:Genetics:  
A:Gene: MYPU\_1970  
A:Genetic code: SGCG  
C:Superfamily: Ureaplasma urealyticum hypothetical protein UUS92

Query Match 58.8%; Score 40; DB 2; Length 606;  
Best Local Similarity 50.0%; Pred. No. 65;  
Matches 7; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 1 KKAVALLPVLLAA 14  
:|:|||||  
Db 32 RKSSILPVLVA 45

RESULT 14  
MMHUND  
nidogen precursor - human  
N:Alternate names: entactin  
C:Species: Homo sapiens (man)  
C:Date: 30-Sep-1991 #sequence\_revision 30-Sep-1991 #text\_change 09-Jul-2004  
C:Accession: A33322; A32437; A61367  
R:Nagayoshi, T.; Sanborn, D.; Hickok, N.J.; Olsen, D.R.; Pazio, M.J.; Chu, M.L.; Knowl  
DNA 8, 581-594, 1989  
A:Title: Human nidogen: complete amino acid sequence and structural domains deduced from  
A:Reference number: A33322; MUID:90091745; PMID:2574658  
A:Accession: A33322  
A:Molecule type: mRNA  
A:Residues: 1-1247 <NAG>  
A:Cross-references: UNIPROT:P14543; EMBL:M30269  
R:Olsen, D.R.; Nagayoshi, T.; Pazio, M.; Mattei, M.G.; Passage, E.; Weil, D.; Timpl, R.,  
Am. J. Hum. Genet. 44, 876-885, 1989  
A:Title: Human nidogen: cDNA cloning, cellular expression, and mapping of the gene to ch  
A:Reference number: A32437; MUID:89270475; PMID:2471408  
A:Accession: A32437  
A:Molecule type: mRNA  
A:Residues: 667-1247 <OLG>  
A:Cross-references: EMBL:M27445; NID:G602466; PIDN:AAAS7261.1; PID:G602467  
R:Fazio, M.J.; O'Leary, J.; Kaehaeri, V.M.; Chen, Y.Q.; Saitta, B.; Uitto, J.  
J. Invest. Dermatol. 97, 281-285, 1991  
A:Title: Human nidogen gene: structural and functional characterization of the 5'-flanki  
A:Reference number: A61367; MUID:91302882; PMID:1906509  
A:Accession: A61367  
A:Molecule type: DNA  
A:Residues: 1-28 <FAZ>  
C:Comment: This protein is a basement membrane glycoprotein that forms a complex with la  
C:Genetics:  
A:Gene: GDB:NID  
A:Cross-references: GDB:120236; OMIM:131390  
A:Map position: 1q43-1q43  
C:Superfamily: nidogen; EGF homology; LDL receptor YWTD-containing repeat homology; thyr  
C:Keywords: basement membrane; beta-hydroxyasparagine; calcium binding; cell binding; co  
protein

F;1-28/Domain: signal sequence #status predicted <SIG>  
F;29-1247/Product: nidogen #status predicted <MAT>  
F;390-425/Domain: EGF homology <EG1>  
F;672-708/Domain: EGF homology <EG2>  
F;702-704/Region: cell attachment (R-G-D) motif  
F;714-750/Domain: EGF homology <EG3>  
F;762-800/Domain: EGF homology <EG4>  
F;806-839/Domain: EGF homology <EG5>  
F;849-919/Domain: thyroglobulin type I repeat homology <THV1>  
F;990-1032/Domain: LDL receptor YWTD-containing repeat homology <YW1>  
F;1033-1075/Domain: LDL receptor YWTD-containing repeat homology <YW2>  
F;1076-1120/Domain: LDL receptor YWTD-containing repeat homology <YW3>  
F;1121-1160/Domain: LDL receptor YWTD-containing repeat homology <YW4>  
F;1161-1197/Domain: LDL receptor YWTD-containing repeat homology <YW5>  
F;1212-1243/Domain: EGF homology <EG6>  
F;289,296/Binding site: sulfate (Tyr) (covalent) #status predicted  
F;729,819/Modified site: erythro-beta-hydroxyasparagine (Asn) #status predicted  
F;756/Cross-link: isopeptide (Gln) (interchain to Lys N6-amino of laminin) #status predi  
F;1137/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 58.8%; Score 40; DB 1; Length 1247;  
 Best Local Similarity 72.7%; Pred. No. 1.3e+02;  
 Matches 8; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 5 AVLLPVLAAAP 15  
 :|:|:|:|:|  
 Db 15 ALLPLLLAGP 25

RESULT 15  
 B44102  
 di-N-acetylchitinase (EC 3.2.1.1) - bovine (fragment)  
 N:Alternate names: chitinase; lysosomal glycosidase  
 C:Species: Bos primigenius taurus (cattle)  
 C:Date: 27-Apr-1993 #sequence\_revision 18-Nov-1994 #text\_change 18-Mar-1997  
 C:Accession: B44102  
 R:Fisher, K.J.; Aronson Jr., N.N.  
 J. Biol. Chem. 267, 19607-19616, 1992  
 A:Title: Cloning and expression of the cDNA sequence encoding the lysosomal glycosidase  
 A:Reference number: A44102; MUID:92406917; PMID:1527079  
 A:Accession: B44102  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-175 <FIS>  
 A:Note: sequence inconsistent with the nucleotide translation  
 A:Note: sequence extracted from NCBI backbone (NCBI:113986)  
 C:Keywords: glycosidase; hydrolase

Query Match 57.4%; Score 39; DB 2; Length 175;  
 Best Local Similarity 57.1%; Pred. No. 30;  
 Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 2 KAAAVLLPVLAAAP 15  
 :|:|:|:|:|  
 Db 21 RLAPLLPVLAAAP 34

RESULT 16  
 E84187  
 hypothetical protein Vng0274c [imported] - Halobacterium sp. NRC-1  
 C:Species: Halobacterium sp. NRC-1  
 C:Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 09-Jul-2004  
 R:Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Laeky, S.; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablon  
 Jung, K.H.; Alam, M.; Freitas, T.  
 Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000  
 A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li  
 A:Title: Genome sequence of Halobacterium species NRC-1.  
 A:Reference number: A84160; MUID:20504483; PMID:11016950  
 A:Accession: E84187  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-201 <STO>  
 A:Cross-references: UNIPROT:Q9HSE2; GB:AE004437; NID:g10579909; PIDN:AAG18865.1; GSPDB:G  
 A:Gene: VNG0274C

Query Match 57.4%; Score 39; DB 2; Length 201;  
 Best Local Similarity 69.2%; Pred. No. 34;  
 Matches 9; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 2 KAAAVLLPVLAA 14  
 :|:|:|:|:|  
 Db 19 RAAAVLPVVDAA 31

RESULT 17  
 T47186  
 hypothetical protein DKFZp434B1272.1 - human (fragment)  
 C:Species: Homo sapiens (man)  
 C:Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 09-Jul-2004  
 C:Accession: T47186

R:Poustka, A.; Wellenreuther, R.; Mewes, H.W.; Weil, B.; Wiemann, S.  
 submitted to the Protein Sequence Database, March 2000  
 A:Reference number: Z24374  
 A:Accession: T47186  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-227 <AAA>  
 A:Cross-references: UNIPROT:Q9NSM3; EMBL:AL162032  
 A:Experimental source: adult testis; clone DKFZp434B1272  
 C:Genetics:  
 A:Note: DKFZp434B1272.1

Query Match 57.4%; Score 39; DB 2; Length 227;  
 Best Local Similarity 80.0%; Pred. No. 39;  
 Matches 8; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 KAAAVLLPVL 11  
 :|:|:|:|:|  
 Db 113 KAAVLLPIL 122

RESULT 18  
 H72469  
 hypothetical protein APE2401 - Aeropyrum pernix (strain K1)  
 C:Species: Aeropyrum pernix  
 C:Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 09-Jul-2004  
 C:Accession: H72469  
 R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takai  
 awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.;  
 DNA Res. 6, 83-101, 1999  
 A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropy  
 A:Reference number: A72450; MUID:99310339; PMID:10382966  
 A:Accession: H72469  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-252 <KAW>  
 A:Cross-references: UNIPROT:Q9Y984; DBJ:AP000064; NID:g5105945; PIDN:BAA81416.1; PID:d  
 A:Experimental source: strain K1  
 C:Genetics:  
 A:Gene: APE2401

Query Match 57.4%; Score 39; DB 2; Length 252;  
 Best Local Similarity 53.8%; Pred. No. 42;  
 Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 3 AAAVLLPVLAAAP 15  
 :|:|:|:|:|  
 Db 149 ATALIPVLLSMP 161

RESULT 19  
 B8336  
 hypothetical protein PA2474 [imported] - Pseudomonas aeruginosa (strain PAO1)  
 C:Species: Pseudomonas aeruginosa  
 C:Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 09-Jul-2004  
 C:Accession: B8336  
 R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; B  
 adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,  
 ; Lory, S.; Olson, M.V.  
 Nature 406, 959-964, 2000  
 A:Title: Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic path  
 A:Reference number: A82950; MUID:20437337; PMID:10984043  
 A:Accession: B8336  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-303 <STO>  
 A:Cross-references: UNIPROT:Q9I108; GB:AE004675; GB:AE004091; NID:g9948522; PIDN:AAG058;  
 A:Experimental source: strain PAO1  
 C:Genetics:  
 A:Gene: PA2474

Query Match 57.4%; Score 39; DB 2; Length 303;  
 Best Local Similarity 66.7%; Pred. No. 50;

Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 4 AAALLPVLAAAP 15  
| : : : : |  
Db 96 AEILFVLLADP 107

RESULT 20  
AC0389  
phosphate binding protein [imported] - Yersinia pestis (strain CO92)  
C:Species: Yersinia pestis  
C:Date: 02-Nov-2001 #sequence\_revision 02-Nov-2001 #text\_change 09-Jul-2004  
C:Accession: AC0389  
R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.; deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.; il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell, Nature 413, 523-527, 2001  
A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.  
A:Reference number: AB0001; MUID:21470413; PMID:11586360  
A:Accession: AC0389  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-354 <KUR>  
A:Cross-references: UNIPROT:Q8ZC23; GB:AL590842; PIDN:CAC92438.1; PID:g15981140; GSPDB:G  
C:Genetics:  
A:Gene: pstS

Query Match 57.4%; Score 39; DB 2; Length 354;  
Best Local Similarity 50.0%; Pred. No. 58;  
Matches 7; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 2 KAAAVLLPVLAAAP 15  
| : : : : |  
Db 40 RVALVLPVTLASP 53

RESULT 21  
G84971  
ompF-like porin [imported] - Buchnera sp. (strain APS)  
C:Species: Buchnera sp.  
C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 23-Mar-2001  
C:Accession: G84971  
R:Shigenobu, S.; Watanabe, H.; Hattori, M.; Sakaki, Y.; Ishikawa, H.  
Nature 407, 81-86, 2000  
A:Title: Genome sequence of the endocellular bacterial symbiont of aphids Buchnera sp. A  
A:Reference number: AB4930; MUID:20445173; PMID:10993077  
A:Accession: G84971  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-382 <STO>  
A:Cross-references: GB:AP000398; GSPDB:GN00144  
A:Experimental source: strain APS  
C:Genetics:  
A:Gene: ompF; BU359  
C:Superfamily: outer membrane protein phoE

Query Match 57.4%; Score 39; DB 2; Length 382;  
Best Local Similarity 50.0%; Pred. No. 62;  
Matches 7; Conservative 6; Mismatches 1; Indels 0; Gaps 0;

QY 1 KKAARVLLPVLAA 14  
| : : : : |  
Db 4 RKSLAWVPMLLAA 17

RESULT 22  
F72554  
hypothetical protein APE1722 - Aeropyrum pernix (strain K1)  
C:Species: Aeropyrum pernix  
C:Date: 20-Aug-1999 #sequence\_revision 20-Aug-1999 #text\_change 09-Jul-2004  
C:Accession: F72554  
R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takah  
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K

DNA Res. 6, 83-101, 1999  
A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr  
A:Reference number: A72450; MUID:99310339; PMID:10382966  
A:Accession: F72554  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-389 <KAW>  
A:Cross-references: UNIPROT:Q9VB75; DDBJ:AP000062; NID:g5105244; PIDN:BAA80723.1; PID:d:1  
A:Experimental source: strain K1  
C:Genetics:  
A:Gene: APE1722

Query Match 57.4%; Score 39; DB 2; Length 389;  
Best Local Similarity 72.7%; Pred. No. 63;  
Matches 8; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 KKAARVLLPVL 11  
| : : : : |  
Db 294 KKTAKILLPVL 304

RESULT 23  
T35263  
probable carboxypeptidase - Streptomyces coelicolor  
C:Species: Streptomyces coelicolor  
C:Date: 05-Nov-1999 #sequence\_revision 05-Nov-1999 #text\_change 09-Jul-2004  
C:Accession: T35263  
R:Oliver, K.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.  
submitted to the EMBL Data Library, April 1999  
A:Reference number: Z21573  
A:Accession: T35263  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-432 <OLI>  
A:Cross-references: UNIPROT:Q9X7P4; EMBL:AL049587; PIDN:CAB40688.1; GSPDB:GN00070; SCORE  
A:Experimental source: strain A3(2)  
C:Genetics:  
A:Gene: SCOEDB:SCSF2A.21

Query Match 57.4%; Score 39; DB 2; Length 432;  
Best Local Similarity 61.5%; Pred. No. 70;  
Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 3 AAALLPVLAAAP 15  
| : : : : |  
Db 15 AASLVFSLTAP 27

RESULT 24  
AB3032  
hypothetical protein Atu3863 [imported] - Agrobacterium tumefaciens (strain C58, Dupont)  
C:Species: Agrobacterium tumefaciens  
C:Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 09-Jul-2004  
C:Accession: AB3032  
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L  
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClell  
; Karp, P.; Romero, P.; Zhang, S.  
Science 294, 2317-2323, 2001  
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,  
ster, E.W.  
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.  
A:Reference number: AB2577; MUID:21608550; PMID:11743193  
A:Accession: AB3032  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-493 <KUR>  
A:Cross-references: UNIPROT:Q8U966; GB:AE008689; PIDN:AAL44672.1; PID:g17742298; GSPDB:G  
A:Experimental source: strain C58 (Dupont)  
C:Genetics:  
A:Gene: Atu3863  
A:Map position: linear chromosome

Query Match 57.4%; Score 39; DB 2; Length 493;

```

Best Local Similarity 63.6%; Pred. No. 79;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 5 AVLLPVLAAAP 15
   ||:|:|:|
Db 190 AVVAFILIAAP 200

RESULT 25
F90031
PTS system, arbutin-like IIBC component [imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C>Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C:Accession: F90031
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A:Reference number: A89758; MUID:21311952; PMID:11418146
A:Accession: F90031
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-534 <KUR>
A:Cross-references: UNIPROT:Q99RV0; GB:BA000018; PID:gl3702123; PIDN:BAB43415.1; GSPDB:C
A:Experimental source: strain N315
C:Genetics:
A:Gene: glvC
C:Superfamily: phosphotransferase system glucose-specific enzyme II, factor II; phosphot
Query Match 57.4%; Score 39; DB 2; Length 534;
Best Local Similarity 57.1%; Pred. No. 85;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 KKAADVLLPVLAA 14
   ||:|:|:|
Db 302 KKVAAALLPATLTA 315

RESULT 26
T51035
hypothetical protein B15120.50 [imported] - Neurospora crassa
C:Species: Neurospora crassa
C>Date: 21-Jul-2000 #sequence_revision 21-Jul-2000 #text_change 09-Jul-2004
C:Accession: T51035
R:Schulte, U.; Aign, V.; Hoheisel, J.; Brandt, P.; Partmann, B.; Holland, R.; Nyakatura,
submitted to the Protein Sequence Database, July 2000
A:Reference number: Z25286
A:Accession: T51035
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-548 <SCH>
A:Cross-references: UNIPROT:Q9P3K6; EMBL:AL389900; GSPDB:GN00116; NCSP:B15120.50
A:Experimental source: BAC clone B15120; strain OR7A
C:Genetics:
A:Gene: NCSP:B15120.50
A:Map position: 6

Query Match 57.4%; Score 39; DB 2; Length 548;
Best Local Similarity 61.5%; Pred. No. 87;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 KKAADVLLPVLAA 13
   ||:|:|:|
Db 469 KKGVAVLQPIVLA 481

RESULT 27
A96254
exsB protein (AJ225561) [imported] - Agrobacterium tumefaciens (strain C58, Cereon)
C:Species: Agrobacterium tumefaciens
C>Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 09-Jul-2004
C:Accession: A96254

R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.
Science 294, 2323-2328, 2001
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tu
A:Reference number: A97359; MUID:21608551; PMID:11743194
A:Accession: A96254
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-602 <KUR>
A:Cross-references: UNIPROT:Q8U966; GB:AE007870; PIDN:AAK89555.1; PID:gl5159439; GSPDB:C
C:Genetics:
A:Gene: AGR_L 1966
A:Map position: linear chromosome

Query Match 57.4%; Score 39; DB 2; Length 602;
Best Local Similarity 63.6%; Pred. No. 95;
Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 5 AVLLPVLAAAP 15
   ||:|:|:|
Db 299 AVVAFILIAAP 309

RESULT 28
TIZM1
trypsin/factor XIIa inhibitor precursor - maize
N:Alternate names: Hageman factor inhibitor
C:Species: Zea mays (maize)
C>Date: 28-Aug-1985 #sequence_revision 31-Dec-1992 #text_change 09-Jul-2004
C:Accession: S20850; A01327
R:Wen, L.; Huang, J.K.; Zen, K.C.; Johnson, B.H.; Muthukrishnan, S.; MacKay, V.; Manney,
Plant Mol. Biol. 18, 813-814, 1992
A:Title: Nucleotide sequence of a cDNA clone that encodes the maize inhibitor of trypsin
A:Reference number: S20850; MUID:92216060; PMID:1558956
A:Accession: S20850
A:Molecule type: mRNA
A:Residues: 1-155 <WEN>
A:Cross-references: UNIPROT:P01088; EMBL:X54064; NID:q22326; PIDN:CAA37998.1; PID:q2232
R:Mahoney, W.C.; Hermodson, M.A.; Jones, B.; Powers, D.D.; Corfman, R.S.; Reeck, G.R.
J. Biol. Chem. 259, 8412-8416, 1984
A:Title: Amino acid sequence and secondary structural analysis of the corn inhibitor of
A:Reference number: A01327; MUID:84239823; PMID:6610678
A:Accession: A01327
A:Molecule type: protein
A:Residues: 29-48, 'C', 50-55, 'R', 57-60, 'PR', 61-106, 'A', 108-116, 'A', 118-119, 'Q', 121-134,
A>Note: 117-Glu was also found
C:Comment: Five disulfide bonds are present.
C:Superfamily: wheat alpha-amylase inhibitor
C:Keywords: disulfide bond; serine proteinase inhibitor
F:1-28/Domain: signal sequence #status predicted <SIG>
F:29-138/Product: trypsin/factor XIIa inhibitor #status experimental <MAT>
F:139-155/Domain: carboxyl-terminal propeptide #status predicted <CTP>
F:62/inhibitory site: Arg (trypsin) #status experimental

Query Match 55.9%; Score 38; DB 1; Length 155;
Best Local Similarity 83.3%; Pred. No. 40;
Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 3 AAAVLLPVLAA 14
   ||:|:|:|
Db 15 AAAVLLSVLAA 26

RESULT 29
S46667
meiosis-specific protein REC104 - yeast (Saccharomyces cerevisiae)
N:Alternate names: protein YHR157w
C:Species: Saccharomyces cerevisiae
C>Date: 03-May-1994 #sequence_revision 27-Jan-1995 #text_change 09-Jul-2004
C:Accession: S46667; A48424; S46753
R:Hollingsworth, N.M.; Johnson, A.D.
Genetics 133, 785-797, 1993
A:Title: A conditional allele of the Saccharomyces cerevisiae HOP1 gene is suppressed b;

```



A:Reference number: S46667; MUID:93316080; PMID:8462842

A:Accession: S46667

A:Molecule type: DNA

A:Residues: 1-182 <HOL>

A:Cross-references: UNIPROT:P33323; EMBL:S58278; NID:G299060; PIDN:AAB26085.1; PID:G299060

R:Galbraith, A.M.; Malone, R.E.

Dev. Genet. 13, 392-402, 1992

A:Title: Characterization of REC104, a gene required for early meiotic recombination in

A:Reference number: A48424; MUID:93278888; PMID:1304422

A:Accession: A48424

A:Molecule type: DNA

A:Residues: 1-182 <GAL>

A:Cross-references: EMBL:T15007

A:Note: sequence extracted from NCBI backbone (NCBIN:133302, NCBI:P133395)

R:Macri, C.

submitted to the EMBL Data Library, June 1994

A:Description: The sequence of *S. cerevisiae* cosmid 9666.

A:Reference number: S46752

A:Accession: S46753

A:Molecule type: DNA

A:Residues: 1-182 <MAC>

A:Cross-references: EMBL:U10397; NID:G500647; PID:G500649; MIPS:YHR157W

C:Genetics:

A:Gene: SGD:REC104

A:Cross-references: SGD:S0001200; MIPS:YHR157W

A:Map position: 8R

C:Function:

A:Description: required for early meiotic recombination

C:Superfamily: Saccharomyces cerevisiae meiosis-specific protein REC104

C:Keywords: meiosis

Query Match 55.9%; Score 38; DB 2; Length 182;

Best Local Similarity 53.3%; Pred. No. 46;

Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy 1 KKAVALLPVLLAAP 15

||:|:|:|:|

Db 78 KKSATVSLPFLIGDP 92

RESULT 30

T25330

hypothetical protein T26H5.1 - *Caenorhabditis elegans*

C:Species: *Caenorhabditis elegans*

C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999

C:Accession: T25330

R:Gardner, A.

submitted to the EMBL Data Library, November 1996

A:Reference number: Z20017

A:Accession: T25330

A:Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-246 <WIL>

A:Cross-references: EMBL:Z82056; PIDN:CAB04853.1; GSPDB:GN00023; CESP:T26H5.1

C:Genetics:

A:Gene: CESP:T26H5.1

A:Map position: 5

A:Introns: 102/1; 163/3

Query Match

Best Local Similarity 55.9%; Score 38; DB 2; Length 246;

Matches 6; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 3 AAVALLPVLLAAP 15

|:|:|:|:|

Db 140 AVTFMIPILAAAP 152

RESULT 31

T35589

probable secreted protein - *Streptomyces coelicolor*

C:Species: *Streptomyces coelicolor*

C:Date: 05-Nov-1999 #sequence\_revision 05-Nov-1999 #text\_change 09-Jul-2004

A:Accession: T35589; S37564

R:Saunders, D.; Harris, D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.

submitted to the EMBL Data Library, August 1998

A:Reference number: Z21583

A:Accession: T35589

A:Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-249 <SAU>

A:Cross-references: UNIPROT:P40179; EMBL:AL031317; PIDN:CAA20416.1; GSPDB:GN00070; SCOE

A:Experimental source: strain A3(2)

R:Duchene, A.M.; Kieser, K.; Hopwood, D.; Thompson, C.; Mazodier, P.

submitted to the EMBL Data Library, September 1993

A:Description: Molecular characterization of two *groEL* genes in *Streptomyces coelicolor*

A:Reference number: S37564

A:Accession: S37564

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 121-249 <DUC>

A:Cross-references: EMBL:X75206; NID:G406595; PIDN:CAA53017.1; PID:G406596

C:Genetics:

A:Gene: SCOE:DB:SC6G4.38

Query Match 55.9%; Score 38; DB 2; Length 249;

Best Local Similarity 66.7%; Pred. No. 62;

Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 4 RAVALLPVLLAAP 15

||:|:|:|:|

Db 14 AAALLPVLALLP 25

RESULT 32

T03239

probable chitinase (EC 3.2.1.14) precursor - rice

C:Species: *Oryza sativa* (rice)

C:Date: 24-Mar-1999 #sequence\_revision 24-Mar-1999 #text\_change 09-Jul-2004

C:Accession: T03239

R:Anuratha, C.S.; Mew, T.; Muthukrishnan, S.

submitted to the EMBL Data Library, September 1993

A:Description: Induction of chitinases and beta-glucanases in *Rhizoctonia solani* infecte

A:Reference number: Z14854

A:Accession: T03239

A:Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: mRNA

A:Residues: 1-335 <ANU>

A:Cross-references: UNIPROT:Q42970; EMBL:U02286; NID:G495302; PIDN:AAA18585.1; PID:G4953

A:Experimental source: leaf, strain IR58

C:Superfamily: lectin-related plant chitinase; hevein chitin-binding domain homology; p

C:Keywords: glycosidase; hydrolase; polysaccharide degradation

F:1-32/Domain: signal sequence #status predicted <SIG>

F:33-335/Product: chitinase #status predicted <MAT>

F:33-74/Domain: hevein chitin-binding domain homology <HCB>

F:90-325/Domain: plant chitinase homology <PCH>

Query Match 55.9%; Score 38; DB 2; Length 335;

Best Local Similarity 71.4%; Pred. No. 81;

Matches 10; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 KKAVALLPVLLAA 14

|||||:|:|:|

Db 12 KKAALVALVLA 25

RESULT 33

S39979

chitinase (EC 3.2.1.14) - rice

C:Species: *Oryza sativa* (rice)

C:Date: 18-Feb-1994 #sequence\_revision 03-Aug-1995 #text\_change 09-Jul-2004

C:Accession: S39979

R:Nishizawa, Y.; Kishimoto, N.; Saito, A.; Hibi, T.

Mol. Gen. Genet. 241, 1-10, 1993

A:Title: Sequence variation, differential expression and chromosomal location of rice ch

A:Reference number: S39979; MUID:94049667; PMID:7901749

A:Accession: S39979

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-339 <NIS>

A:Cross-references: UNIPROT:Q43294; EMBL:X56787

C:Superfamily: lectin-related plant chitinase; hevein chitin-binding domain homology; pl

C:Keywords: glycosidase; hydrolase; polysaccharide degradation

F:33-74/Domain: hevein chitin-binding domain homology <HCB>

F:91-329/Domain: plant chitinase homology <PCH>

Query Match 55.9%; Score 38; DB 2; Length 339;  
Best Local Similarity 71.4%; Pred. No. 82;  
Matches 10; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 KKA0AVLLPVLLAA 14

DB 12 KKA0AVALLVLA 25

RESULT 34

S40414

C:Species: Oryza sativa (rice)

C:Date: 25-Dec-1994 #sequence\_revision 03-Aug-1995 #text\_change 09-Jul-2004

C:Accession: S40414

R:Nishizawa, Y.

submitted to the EMBL Data Library, November 1990

A:Reference number: S40414

A:Accession: S40414

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-340 <NIS>

A:Cross-references: UNIPROT:Q43294; EMBL:X56787; NID:G407471; PID:G4074

C:Superfamily: lectin-related plant chitinase; hevein chitin-binding domain homology; pl

C:Keywords: glycosidase; hydrolase; polysaccharide degradation

F:33-74/Domain: hevein chitin-binding domain homology <HCB>

F:92-330/Domain: plant chitinase homology <PCH>

Query Match 55.9%; Score 38; DB 2; Length 340;  
Best Local Similarity 71.4%; Pred. No. 82;  
Matches 10; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 KKA0AVLLPVLLAA 14

DB 12 KKA0AVALLVLA 25

RESULT 35

D64888

C:Species: Escherichia coli

C:Date: 12-Sep-1997 #sequence\_revision 17-Sep-1997 #text\_change 09-Jul-2004

C:Accession: D64888

R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co

A.; Rose, D.J.; Mau, B.; Shao, Y.

Science 277, 1453-1462, 1997

A:Title: The complete genome sequence of Escherichia coli K-12.

A:Reference number: A64720; MUID:97426617; PMID:9278503

A:Accession: D64888

A>Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-377 <BIAT>

A:Cross-references: UNIPROT:P77747; GB:AE000234; GB:U00096; NID:G1787633; PIDN:AAC74459.

C:Superfamily: outer membrane protein phoE

C:Keywords: membrane protein; porin; trimmer

F:1-21/Domain: signal sequence #status predicted <SIG>

F:22-377/Product: probable outer membrane porin blp377 #status predicted <OPP>

Query Match 55.9%; Score 38; DB 2; Length 377;  
Best Local Similarity 61.5%; Pred. No. 90;  
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 KAAAVLLPVLLAA 14

DB 4 KVALLLPALLAA 16

RESULT 36

JC4802

C:Species: Thermoactinomyces sp.

A:Variety: strain E79

C:Date: 15-Aug-1996 #sequence\_revision 15-Oct-1996 #text\_change 09-Jul-2004

C:Accession: JC4802

R:Lee, J.K.; Kim, Y.O.; Kim, H.K.; Park, Y.S.; Oh, T.K.

Biosci. Biotechnol. Biochem. 60, 840-846, 1996

A:Title: Purification and characterization of a thermostable alkaline protease from The

A:Reference number: JC4802; MUID:96261070; PMID:8704314

A:Accession: JC4802

A:Molecule type: DNA

A:Residues: 1-384 <LEE>

A:Cross-references: UNIPROT:Q56365; GB:U31759; NID:G1389689; PIDN:AAB36499.1; PID:G1683

A:Experimental source: strain E79

C:Comment: This protein is thermostable.

C:Function:

A:Description: extracellular alkaline serine proteinase [validated, MUID:96261070]

C:Superfamily: subtilisin; subtilisin homology

C:Keywords: hydrolase; serine proteinase

F:1-25/Domain: signal sequence #status predicted <SIG>

F:26-106/Domain: propeptide #status predicted <PRO>

F:107-383/Product: alkaline proteinase #status experimental <MAT>

F:134-344/Domain: subtilisin homology <SBT>

F:143,176,330/Active site: Asp, His, Ser #status predicted

Query Match 55.9%; Score 38; DB 2; Length 384;

Best Local Similarity 69.2%; Pred. No. 92;

Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 AAVLLPVLLAAP 15

DB 8 ATALLLVLLAVP 20

RESULT 37

C82614

C:Species: Xylella fastidiosa

C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 09-Jul-2004

C:Accession: C82614

R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen

Nature 406, 151-157, 2000

A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.

A:Reference number: A82515; MUID:20365717; PMID:10910347

A>Note: for a complete list of authors see reference number A59328 below

A:Accession: C82614

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-398 <SIM>

A:Cross-references: UNIPROT:Q9PB23; GB:AE004018; GB:AE003849; NID:G9107093; PIDN:AAF847

A:Experimental source: strain 9a5c

R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; i

Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carriaro, D.M.; Carreir, i

as-Neto, E.; Docena, C.; El-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S.

submitted to GenBank, June 2000

A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Praga, J.S.; Franca, S.C.; Franco, M.C.; Froh

J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.B.; Laig

chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins,

A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, B.C.; Miyaki, C.Y.

P.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.V.; de Sa, R.G.; Santelli,

Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Silva, R.V.; de Silva Jr., W.A.; da Silveir

A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir

M.; Teuhako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; i

A:Reference number: A59328

A:Contents: annotation

## C:Genetics:

A:Gene: XF1992

C:Superfamily: Azospirillum cyclohydrolase II; cyclohydrolase homology

Query Match 55.9%; Score 38; DB 2; Length 398;

Best Local Similarity 63.6%; Pred. No. 95;

Matches 7; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Oy 5 AALLPVLLAAP 15

|:|:|:|:|:|

Db 131 AALLPAIVAAP 141

## RESULT 38

T35476

probable regulatory protein - Streptomyces coelicolor

C:Species: Streptomyces coelicolor

C:Date: 05-Nov-1999 #sequence\_revision 05-Nov-1999 #text\_change 09-Jul-2004

C:Accession: T35476

R:Oliver, K.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.

submitted to the EMBL Data Library, December 1998

A:Reference number: 221579

A:Accession: T35476

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-569 &lt;OLI&gt;

A:Cross-references: UNIPROT:Q92BP0; EMBL:AL034492; PIDN:CAA22503.1; GSPDB:GN00070; SCOE

A:Experimental source: strain A3(2)

C:Genetics:

A:Gene: SCOEDB:SC6C5.05C

Query Match

Best Local Similarity 55.9%; Score 38; DB 2; Length 569;

Matches 10; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Oy 3 AAALLPVLLAAP 15

|:|:|:|:|:|

Db 322 AAPVLLPGLGAAP 334

## RESULT 39

E95296

probable ABC-type iron transport system protein SMA0525 [imported] - Sinorhizobium meliloti

C:Species: Sinorhizobium meliloti

C:Date: 24-Aug-2001 #sequence\_revision 24-Aug-2001 #text\_change 09-Jul-2004

C:Accession: E95296

R:Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; Bows

; Kalman, S.; Keating, D.H.; Palm, C.; Peck, M.C.; Surzycki, R.; Wells, D.H.; Yeh, K.C.

Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001

A:Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium meliloti

A:Reference number: A95262; MUID:21396509; PMID:11481432

A:Accession: E95296

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-601 &lt;KUR&gt;

A:Cross-references: UNIPROT:Q930C1; GB:AE006469; PIDN:AAK64935.1; PID:gl4523357; GSPDB:C

A:Experimental source: strain 1021, megaplasmid pSymA

R:Galibert, F.; Finan, T.M.; Long, S.R.; Fuhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,

pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;

L.; Hyman, R.W.; Jones, T.

Science 293, 668-672, 2001

A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,

heault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.

A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.

A:Reference number: A96039; MUID:21368234; PMID:11474104

A:Contents: annotation

C:Genetics:

A:Gene: SMA0525

A:Genome: plasmid

Query Match

Best Local Similarity 55.9%; Score 38; DB 2; Length 601;

Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Oy 1 KKAVALLPVLLAAP 15

|:|:|:|:|:|

Db 42 KAAVLVLLTALVAAP 56

## RESULT 40

A90644

probable membrane protein [imported] - Escherichia coli (strain O157:H7, substrain RIMD

C:Species: Escherichia coli

C:Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 09-Jul-2004

C:Accession: A90644

R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.

gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shingawa, H.

DNA Res. 8, 11-22, 2001

A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and geno

A:Reference number: A99629; MUID:21156231; PMID:11258796

A:Accession: A90644

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-617 &lt;HAV&gt;

A:Cross-references: UNIPROT:Q8X959; GB:BA000007; PIDN:BA833544.1; PID:gl13359577; GSPDB:G

A:Experimental source: strain O157:H7, substrain RIMD 0509952

C:Genetics:

A:Gene: EC80121

C:Superfamily: Escherichia coli yach protein

Query Match

Best Local Similarity 55.9%; Score 38; DB 2; Length 617;

Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Oy 3 AAVALLPVLLAAP 15

|:|:|:|:|:|

Db 50 AAATVLPATVSAP 62

## RESULT 41

A85495

probable membrane protein yach [imported] - Escherichia coli (strain O157:H7, substrain

C:Species: Escherichia coli

C:Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 09-Jul-2004

C:Accession: A85495

R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew

iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,

Nature 409, 529-533, 2001

A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.

A:Reference number: A85480; MUID:21074935; PMID:11206551

A:Accession: A85495

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-617 &lt;STO&gt;

A:Cross-references: UNIPROT:Q8X959; GB:AE005174; NID:gl2512829; PIDN:AA654421.1; GSPDB:G

A:Experimental source: strain O157:H7, substrain EDL933

C:Genetics:

A:Gene: yach

C:Superfamily: Escherichia coli yach protein

Query Match

Best Local Similarity 55.9%; Score 38; DB 2; Length 617;

Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Oy 3 AAVALLPVLLAAP 15

|:|:|:|:|:|

Db 50 AAATVLPATVSAP 62

## RESULT 42

T31653

hypothetical protein Y57A10A.ff - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 09-Jul-2004

A:Accession: T31653

R:Smyle, R.

submitted to the EMBL Data Library, September 1999

A:Reference number: Z21048  
 A:Accession: T31653  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-1048 <WIL>  
 A:Cross-references: UNIPROT:Q9NA71; EMBL:AL117195; PIDN:CAB55036.1; CESP:Y  
 A:Experimental source: Clone Y57A10A  
 C:Genetics:  
 A:Gene: CESP:Y57A10A.ff  
 A:Introns: 34/1; 82/3; 156/3; 398/3; 556/3; 798/1; 863/3; 1024/1

Query Match 55.9%; Score 38; DB 2; Length 1048;

Best Local Similarity 69.2%; Pred. No. 2.3e+02;

Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 AAAVLLPVLLAAP 15

||| :|||

Db 200 AAAPSPILLAAP 212

RESULT 43

T36946

probable cation-transporting ATPase - Streptomyces coelicolor

C:Species: Streptomyces coelicolor

C>Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004

A:Accession: T36946

R:Seeger, K.J.; Harris, D.; Thomson, N.R.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.  
 submitted to the EMBL Data Library, September 1999

A:Reference number: Z21607

A:Accession: T36946

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-776 <SEE>

A:Cross-references: UNIPROT:Q9RJ01; EMBL:AL109962; PIDN:CAB53131.1; GSPDB:GN00070; SCOE

A:Experimental source: strain A3(2)

C:Genetics:

A:Gene: SCOE:SCJ1.13

F:442-585/Domain: ATPase nucleotide-binding domain homology <ATN>

Query Match 55.1%; Score 37.5; DB 2; Length 776;

Best Local Similarity 52.6%; Pred. No. 2.1e+02;

Matches 10; Conservative 3; Mismatches 1; Indels 5; Gaps 1;

QY 2 KAAAVL-----PVLLAAP 15

:||| :|||

Db 270 RAVAVLVWATPCPLLAAAP 288

RESULT 44

CCDV3M

cytochrome c3 precursor [validated] - Desulfovibrio vulgaris (strain Miyazaki)

C:Species: Desulfovibrio vulgaris

C>Date: 31-Oct-1980 #sequence\_revision 10-Oct-1997 #text\_change 09-Jul-2004

C:Accession: S33874; A00125

R:Kitamura, M.; Ozawa, K.; Kojima, S.; Kumagai, I.; Akutsu, H.; Miura, K.

Protein Seq. Data Anal. 5, 193-196, 1993

A:Title: The primary structure of pre-cytochrome c(3) from Desulfovibrio vulgaris (Miyaz

A:Reference number: S33874

A:Accession: S33874

A:Molecule type: DNA

A:Residues: 1-130 <KIT>

A:Cross-references: UNIPROT:P00132; DBJ:D31702; NID:g496361; PIDN:BAA06511.1; PID:g4963

R:Shinkai, W.; Hase, T.; Yagi, T.; Matsubara, H.

J. Biochem. 87, 1747-1756, 1980

A:Title: Amino acid sequence of cytochrome c-3 from Desulfovibrio vulgaris, Miyazaki.

A:Reference number: A00125; MUID:80249474; PMID:6249799

A:Accession: A00125

A:Molecule type: protein

A:Residues: 24-64, 'N', 66-130 <SHI>

R:Higuchi, Y.; Kusunoki, M.; Matsura, Y.; Yasuoka, N.; Kakudo, M.

J. Mol. Biol. 172, 109-139, 1984

A:Title: Refined structure of cytochrome c-3 at 1.8 angstrom resolution.

A:Reference number: A49705; MUID:84114880; PMID:6319712  
 A:Contents: annotation; X-ray crystallography, 1.8 angstroms  
 R:Higuchi, Y.; Kusunoki, M.; Matsura, Y.; Yasuoka, N.; Kakudo, M.  
 submitted to the Brookhaven Protein Data Bank, November 1983  
 A:Reference number: A50415; PDB:2CDV

A:Contents: annotation; X-ray crystallography, 1.8 angstroms, residues 24-64, 'N', 66-130  
 C:Function:

A:Description: accepts electrons from cytochrome-c3 hydrogenase (EC 1.12.2.1) and transp

A:Pathway: sulfate respiration

C:Superfamily: cytochrome c3; cytochrome c3 homology

C:Keywords: Chromoprotein; electron transfer; heme; iron; metalloprotein; sulfate respi.

F:24-130/Product: cytochrome c3 #status experimental <MAT>

F:43-129/Domain: cytochrome c3 homology <CC3>

F:45-57/Binding site: heme iron (His) (axial ligands) #status experimental

F:48-106/Binding site: heme iron (His) (axial ligands) #status experimental

F:53-56/Binding site: heme (Cys) (covalent) #status experimental

F:58-75/Binding site: heme iron (His) (axial ligands) #status experimental

F:69-74/Binding site: heme (Cys) (covalent) #status experimental

F:93-129/Binding site: heme iron (His) (axial ligands) #status experimental

F:102-105/Binding site: heme (Cys) (covalent) #status experimental

F:123-128/Binding site: heme (Cys) (covalent) #status experimental

Query Match 54.4%; Score 37; DB 1; Length 130;

Best Local Similarity 53.8%; Pred. No. 50;

Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 3 AAAVLLPVLLAAP 15

||| :|||

Db 13 AVATAMPALAAAP 25

RESULT 45

H75531

cytochrome c-type biogenesis protein CcmH - Deinococcus radiodurans (strain R1)

C:Species: Deinococcus radiodurans

C>Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004

C:Accession: H75531

R:White, O.; Easen, J.A.; Heideberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;

M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.;

S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.

Science 286, 1571-1577, 1999

A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.

A:Reference number: A75250; MUID:20036896; PMID:10567266

A:Accession: H75531

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-199 <WHI>

A:Cross-references: UNIPROT:Q9RXG9; GB:AE001894; GB:AE000513; NID:g6458011; PIDN:AAF099.

A:Experimental source: strain R1

C:Genetics:

A:Gene: DR0344

A:Map position: 1

Query Match 54.4%; Score 37; DB 2; Length 199;

Best Local Similarity 60.0%; Pred. No. 74;

Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 KKAAVLLPVLLAAP 15

||| :|||

Db 7 KGAQAVLTALLASP 21

RESULT 46

AC2793

conserved hypothetical protein Atu1762 [Imported] - Agrobacterium tumefaciens (strain C

C:Species: Agrobacterium tumefaciens

C>Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 09-Jul-2004

C:Accession: AC2793

R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, Y.

erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClell

; Karp, P.; Romero, P.; Zhang, S.

Science 294, 2317-2323, 2001

A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,

ster, E.W.  
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.  
A:Reference number: AB2577; MUID:21608550; PMID:11743193  
A:Accession: AC2793  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-254 <KUR>  
A:Cross-references: UNIPROT:Q8UEJ5; GB:AE008698; PIDN:AAL42761.1; PID:g17740203; GSPDB:C  
A:Experimental source: strain C58 (Dupont)  
C:Genetics:  
A:Gene: Atul762  
A:Map position: circular chromosome

Query Match 54.4%; Score 37; DB 2; Length 254;  
Best Local Similarity 66.7%; Pred. No. 92;  
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KKAQAVLLPVLL 12  
|||:|||||  
Db 6 KKAQALLPLLL 17

RESULT 47  
A97572  
hypothetical protein AGR\_C\_3235 [imported] - Agrobacterium tumefaciens (strain C58, Cere  
C:Species: Agrobacterium tumefaciens  
C:Date: 30-Sep-2001 #sequence\_revision 30-Sep-2001 #text\_change 09-Jul-2004  
C:Accession: A97572  
A:; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;  
Science 294, 2323-2328, 2001  
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum  
A:Reference number: A97359; MUID:21608551; PMID:11743194  
A:Accession: A97572  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-254 <KUR>  
A:Cross-references: UNIPROT:Q8UEJ5; GB:AE007869; PIDN:AAK87530.1; PID:g15156862; GSPDB:C  
C:Genetics:  
A:Gene: AGR\_C\_3235  
A:Map position: circular chromosome

Query Match 54.4%; Score 37; DB 2; Length 254;  
Best Local Similarity 66.7%; Pred. No. 92;  
Matches 8; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KKAQAVLLPVLL 12  
|||:|||||  
Db 6 KKAQALLPLLL 17

RESULT 48  
A98098  
hypothetical protein ECs3753 [imported] - Escherichia coli (strain O157:H7, substrain RI  
C:Species: Escherichia coli  
C:Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 09-Jul-2004  
C:Accession: A98098  
R:Hayashi, T.; Makino, K.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.  
gaawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.  
DNA Res. 8, 11-22, 2001  
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genc  
A:Reference number: A99629; MUID:21156231; PMID:11258796  
A:Accession: A98098  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-259 <HAY>  
A:Cross-references: UNIPROT:Q46813; GB:BA000007; PIDN:BA371176.1; PID:g13363225; GSPDB:C  
A:Experimental source: strain O157:H7, substrain RIMD 0509952  
C:Genetics:  
A:Gene: ECs3753

Query Match 54.4%; Score 37; DB 2; Length 259;  
Best Local Similarity 61.5%; Pred. No. 94;

Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KKAQAVLLPVLLA 13  
:::|||||  
Db 110 QQESVLLPVLLA 122

RESULT 49  
E85943  
hypothetical protein Z4219 [imported] - Escherichia coli (strain O157:H7, substrain EDL9  
C:Species: Escherichia coli  
C:Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 09-Jul-2004  
C:Accession: E85943  
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew  
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,  
Nature 409, 529-533, 2001  
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.  
A:Reference number: A85480; MUID:21074935; PMID:11206551  
A:Accession: E85943  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-259 <STO>  
A:Cross-references: UNIPROT:Q46813; GB:AE005174; NID:g12517404; PIDN:AAG58009.1; GSPDB:C  
A:Experimental source: strain O157:H7, substrain EDL933  
C:Genetics:  
A:Gene: Z4219

Query Match 54.4%; Score 37; DB 2; Length 259;  
Best Local Similarity 61.5%; Pred. No. 94;  
Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KKAQAVLLPVLLA 13  
:::|||||  
Db 110 QQESVLLPVLLA 122

RESULT 50  
H65071  
hypothetical protein b2880 - Escherichia coli (strain K-12)  
C:Species: Escherichia coli  
C:Date: 12-Sep-1997 #sequence\_revision 17-Sep-1997 #text\_change 09-Jul-2004  
C:Accession: H65071  
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co  
A.; Rose, D.J.; Mau, B.; Shao, Y.  
Science 277, 1453-1462, 1997  
A:Title: The complete genome sequence of Escherichia coli K-12.  
A:Reference number: A64720; MUID:97426617; PMID:9278503  
A:Accession: H65071  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-259 <BLAT>  
A:Cross-references: UNIPROT:Q46813; GB:AE000371; GB:U00096; NID:g1789239; PIDN:AAC75918.  
A:Experimental source: strain K-12, substrain MG1655

Query Match 54.4%; Score 37; DB 2; Length 259;  
Best Local Similarity 61.5%; Pred. No. 94;  
Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KKAQAVLLPVLLA 13  
:::|||||  
Db 110 QQESVLLPVLLA 122

Search completed: June 2, 2005, 01:35:46  
Job time : 23.7188 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 2, 2005, 01:26:26 ; Search time 106.406 Seconds  
(without alignments)  
72.187 Million cell updates/sec

Title: US-10-764-235-2

Perfect score: 68

Sequence: 1 KKAALVLPVLLAAP 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 750 summaries

Database :

UniProt\_03:\*

1: uniprot\_sprot:\*

2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	45	66.2	133	Q6ND81	Q6nd81 rhodopsin
2	45	66.2	356	Q8H8C7	Q8h8c7 oryza sativ
3	44	64.7	497	Q6N2W2	Q6n2w2 rhodopsin
4	43	63.2	190	Q87TJ5	Q87tj5 vibrio para
5	43	63.2	340	Q8FRS4	Q8frs4 corynebacte
6	43	63.2	582	Q99PG1	Q99pg1 cricetus
7	43	63.2	671	Q6AA98	Q6aa98 propionibac
8	42	61.8	177	Q7W3J1	Q7w3j1 bordetella
9	42	61.8	177	Q7W6W6	Q7w6w6 bordetella
10	42	61.8	253	Q73HU0	Q73hu0 wolbachia p
11	42	61.8	260	Q92M69	Q92m69 rhizobium m
12	42	61.8	291	Q98FS7	Q98fs7 rhizobium l
13	42	61.8	293	Q8FYJ9	Q8fyj9 brucella su
14	42	61.8	297	Q8YJ99	Q8yj99 brucella me
15	42	61.8	323	Q7WNE5	Q7wne5 bordetella
16	42	61.8	422	Q89HQ1	Q89hq1 bradyrhizob
17	42	61.8	939	Q49467	Q49467 arabidopsis
18	41	60.3	145	CU65_LOCM1	P45585 locusta mig
19	41	60.3	195	Q7PS92	Q7ps92 anopheles g
20	41	60.3	195	Q8BIH2	Q8bih2 anopheles g
21	41	60.3	222	Q8CPR9	Q8cpr9 kluveromyc
22	41	60.3	231	Q7U481	Q7u481 synecococc
23	41	60.3	659	Q82NQ4	Q82nq4 streptomyce
24	41	60.3	1172	Q9Y8Y4	Q9y8y4 aeropyrum p
25	40	58.8	43	P8BY_SYNEL	Q8dkm3 synecococc
26	40	58.8	67	Q6CX97	Q6cx97 mus musculu
27	40	58.8	117	NUTP_SPIOL	P10976 spinacia ol
28	40	58.8	138	Q854U6	Q854u6 mycobacteri
29	40	58.8	138	Q6W1Z2	Q6w1z2 rhizobium s
30	40	58.8	152	CU64_LOCM1	P11739 locusta mig
31	40	58.8	153	Q92KQ6	Q92kq6 rhizobium m

32	40	58.8	158	2	Q8UDX0	Q8udx0 agrobacteri
33	40	58.8	163	2	Q7CY40	Q7cy40 agrobacteri
34	40	58.8	168	2	Q7S0K8	Q7s0k8 neurospora
35	40	58.8	252	2	Q9HIX9	Q9hix9 thermoplasma
36	40	58.8	279	2	Q89D56	Q89d56 bradyrhizob
37	40	58.8	297	2	Q7QB91	Q7qb91 anopheles g
38	40	58.8	298	2	Q83R77	Q83r77 shigella fl
39	40	58.8	390	2	Q9YDB2	Q9ydb2 aeropyrum p
40	40	58.8	431	2	Q8UI19	Q8ui19 agrobacteri
41	40	58.8	435	2	Q7D1D9	Q7d1d9 agrobacteri
42	40	58.8	450	2	Q70PA4	Q70pa4 melittangiu
43	40	58.8	458	2	Q82DE6	Q82de6 streptomyce
44	40	58.8	483	2	Q6NF46	Q6nf46 corynebacte
45	40	58.8	580	1	MM14_PIG	P50281 homo sapien
46	40	58.8	582	1	MM14_HUMAN	Q9xt90 sus scrofa
47	40	58.8	582	1	MM14_MOUSE	P53690 mus musculu
48	40	58.8	582	1	MM14_RABIT	Q93220 coryctolagus
49	40	58.8	582	1	MM14_RAT	Q10739 rattus norv
50	40	58.8	582	2	Q9GLE4	Q9gle4 bos taurus
51	40	58.8	582	2	Q9XSP0	Q9xsp0 capra hircu
52	40	58.8	582	2	Q6DFU5	Q6dfu5 mus musculu
53	40	58.8	582	2	Q8BXT2	Q8bxt2 mus musculu
54	40	58.8	606	2	Q88R13	Q88r13 mycoplasma
55	40	58.8	639	2	Q9K4H5	Q9k4h5 streptomyce
56	40	58.8	831	2	Q89RR5	Q89rr5 bradyrhizob
57	40	58.8	1114	2	Q86XD7	Q86xd7 homo sapien
58	40	58.8	1247	1	NIDO_HUMAN	P14543 homo sapien
59	39.5	58.1	438	2	Q8KZF3	Q8kzf3 thermus the
60	39.5	58.1	438	2	Q72124	Q72124 thermus the
61	39	57.4	156	2	Q94MP8	Q94mp8 bacterioph
62	39	57.4	175	1	DIAC_BOVIN	Q01458 bos taurus
63	39	57.4	195	2	Q9FC16	Q9fc16 streptomyce
64	39	57.4	201	2	Q9HSE2	Q9hse2 halobacteri
65	39	57.4	210	2	Q6A6R4	Q6a6r4 propionibac
66	39	57.4	227	2	Q9NSM3	Q9nsm3 homo sapien
67	39	57.4	227	2	Q8GRV7	Q8grv7 oryza sativ
68	39	57.4	239	2	Q8XVC4	Q8xvc4 raietonia s
69	39	57.4	250	2	Q8BW80	Q8bw80 mus musculu
70	39	57.4	252	2	Q9Y984	Q9y984 aeropyrum p
71	39	57.4	254	2	Q8BQJ1	Q8bqj1 homo sapien
72	39	57.4	273	2	Q62JC9	Q62jc9 burkholderi
73	39	57.4	273	2	Q63T17	Q63t17 burkholderi
74	39	57.4	295	2	Q8DK28	Q8dk28 pseudomonas
75	39	57.4	303	2	Q91108	Q91108 streptococ
76	39	57.4	305	2	Q8Z299	Q8z299 pyrobaculum
77	39	57.4	354	2	Q66DX4	Q66dx4 yersinia ps
78	39	57.4	354	2	Q82C23	Q82c23 yersinia pe
79	39	57.4	375	2	Q92PC9	Q92pc9 rhizobium m
80	39	57.4	382	1	PORL_BUCAI	P57440 buchnera ap
81	39	57.4	389	2	Q9YB75	Q9yb75 aeropyrum p
82	39	57.4	393	2	Q7Z7M2	Q7z7m2 homo sapien
83	39	57.4	425	2	Q8H4S4	Q8h4s4 oryza sativ
84	39	57.4	432	2	Q9X7P4	Q9x7p4 streptomyce
85	39	57.4	481	2	Q95W90	Q95w90 trypanosoma
86	39	57.4	493	2	Q8U966	Q8u966 agrobacteri
87	39	57.4	498	2	Q7VU08	Q7vu08 bordetella
88	39	57.4	529	2	Q8CNA8	Q8cna8 staphylococ
89	39	57.4	534	2	Q99RV0	Q99rv0 staphylococ
90	39	57.4	534	2	Q7A049	Q7a049 staphylococ
91	39	57.4	534	2	Q7A400	Q7a400 staphylococ
92	39	57.4	534	2	Q6G627	Q6g627 staphylococ
93	39	57.4	534	2	Q6GEB1	Q6geb1 staphylococ
94	39	57.4	548	2	Q9P3K6	Q9p3k6 neurospora
95	39	57.4	560	2	Q6ZMQ1	Q6zmq1 homo sapien
96	39	57.4	578	2	Q6BIL7	Q6bil7 debaryomyce
97	39	57.4	602	2	Q7CTH2	Q7cth2 agrobacteri
98	39	57.4	627	2	Q6WLGA	Q6wlg4 uncultured
99	39	57.4	628	2	Q89IJ7	Q89ij7 bradyrhizob
100	39	57.4	818	2	Q9F3E4	Q9f3e4 streptomyce
101	39	57.4	874	2	Q6QNK2	Q6qnk2 homo sapien
102	39	57.4	882	2	Q8NAQ2	Q8naq2 homo sapien
103	39	57.4	935	2	Q89CC4	Q89cc4 bradyrhizob
104	39	57.4	1245	2	Q8XYE3	Q8xye3 raietonia s

105	39	57.4	1320	2	Q8NDW8	Q8ndw8 homo sapien	178	37	54.4	237	2	Q6NA40	Q6na40 rhodopseudo
106	39	57.4	3970	2	Q93HI8	Q93hi8 streptomyc	179	37	54.4	238	2	Q815M3	Q815m3 bacillus ce
107	38.5	56.6	335	2	Q9VTU8	Q9vtu8 drosophila	180	37	54.4	254	2	Q8UEJ5	Q8uej5 agrobacteri
108	38	55.9	38	1	PSBY_PROMM	P59909 prochloroco	181	37	54.4	255	2	Q7DIE9	Q7die9 rhodopirell
109	38	55.9	111	2	Q9NEI2	Q9ne12 leishmania	182	37	54.4	259	1	YGFN_ECOLI	Y64558 escherichia
110	38	55.9	131	2	Q8XPZ0	Q8xpz0 thermus the	183	37	54.4	259	1	YGFN_ECOLI	P64557 escherichia
111	38	55.9	131	2	Q72JC0	Q72jc0 thermus the	184	37	54.4	259	2	Q8FE83	Q8fe83 escherichia
112	38	55.9	148	2	Q7KSF2	Q7ksf2 drosophila	185	37	54.4	268	2	Q7NVY9	Q7nv99 chromobacte
113	38	55.9	155	1	ITRP_MAIZE	P01088 zea mays (m	186	37	54.4	271	2	Q8XZ16	Q8xz16 raietonia s
114	38	55.9	182	1	R104_YEAST	P33323 saccharomyc	187	37	54.4	283	2	Q7PFC5	Q7pfcs anopheles g
115	38	55.9	185	2	Q946V1	Q946v1 zea mays (m	188	37	54.4	288	2	Q8TEJ0	Q8tej0 homo sapien
116	38	55.9	186	2	Q6Z334	Q6z334 oryza sativ	189	37	54.4	289	1	BLT2_ECOLI	Q69395 escherichia
117	38	55.9	192	2	Q6CB84	Q6cb84 yarrowia li	190	37	54.4	293	2	Q6LFX7	Q6lfx7 photobacter
118	38	55.9	216	2	Q6AVH8	Q6avh8 rattus norv	191	37	54.4	298	2	Q43098	Q43098 psophocarpu
119	38	55.9	219	2	Q7VE95	Q7ve95 prochloroco	192	37	54.4	301	2	Q7NTT2	Q7ntt2 chromobacte
120	38	55.9	240	2	Q7EYB9	Q7eyb9 oryza sativ	193	37	54.4	307	2	Q9AT30	Q9at30 oryza sativ
121	38	55.9	249	1	YBGO_STRCO	P40179 streptomyc	194	37	54.4	310	2	Q54023	Q54023 paracoccus
122	38	55.9	256	2	Q98F68	Q98f68 rhizobium l	195	37	54.4	310	2	Q9L2C8	Q9l2c8 streptomyc
123	38	55.9	278	2	Q89TK8	Q89tk8 bradyrhizob	196	37	54.4	311	2	Q8TIZ7	Q8tiz7 methanosarc
124	38	55.9	286	2	Q6TVY8	Q6tv8 orf virus.	197	37	54.4	313	2	Q7YRG4	Q7yrg4 sus scrofa
125	38	55.9	292	2	Q8C667	Q8c667 mus musculu	198	37	54.4	321	2	Q9BU47	Q9bu47 homo sapien
126	38	55.9	310	2	Q6A7F8	Q6a7f8 propionibac	199	37	54.4	323	2	Q94PC8	Q94pc8 coryphaena
127	38	55.9	323	2	Q8LAF7	Q8laf7 arabidopsis	200	37	54.4	323	2	Q958V0	Q958v0 coryphaena
128	38	55.9	335	2	Q42970	Q42970 oryza sativ	201	37	54.4	323	2	Q958V1	Q958v1 coryphaena
129	38	55.9	337	2	Q9CSD1	Q9csd1 mus musculu	202	37	54.4	323	2	Q958V2	Q958v2 coryphaena
130	38	55.9	340	2	Q43294	Q43294 oryza sativ	203	37	54.4	323	2	Q958V3	Q958v3 coryphaena
131	38	55.9	340	2	Q7DNA1	Q7dna1 oryza sativ	204	37	54.4	323	2	Q958X7	Q958x7 coryphaena
132	38	55.9	344	2	Q7UM64	Q7um64 rhodopirell	205	37	54.4	323	2	Q958X8	Q958x8 coryphaena
133	38	55.9	369	2	Q8RYR7	Q8ryr7 oryza sativ	206	37	54.4	323	2	Q958X9	Q958x9 coryphaena
134	38	55.9	370	2	Q9S7B2	Q9s7b2 zea mays (m	207	37	54.4	323	2	Q958Y0	Q958y0 coryphaena
135	38	55.9	377	1	OMPN_ECOLI	P77747 escherichia	208	37	54.4	323	2	Q958Y1	Q958y1 coryphaena
136	38	55.9	377	2	Q85030	Q85030 escherichia	209	37	54.4	323	2	Q958Y2	Q958y2 coryphaena
137	38	55.9	377	2	Q8CW37	Q8cw37 escherichia	210	37	54.4	334	2	Q6AEG8	Q6aeg8 leifsonia x
138	38	55.9	384	2	Q56365	Q56365 thermoactin	211	37	54.4	335	1	B3G3_CRIGR	Q9wu47 cricetus
139	38	55.9	397	2	Q752P8	Q752p8 ashbya goss	212	37	54.4	335	1	B3G3_MOUSE	P58158 mus musculu
140	38	55.9	398	2	Q87D69	Q87d69 xylella fas	213	37	54.4	335	2	Q62A92	Q62a92 burkholderi
141	38	55.9	398	2	Q9PBZ3	Q9pbz3 xylella fas	214	37	54.4	335	2	Q63NR6	Q63nr6 burkholderi
142	38	55.9	404	2	Q6N7R3	Q6n7r3 rhodopseudo	215	37	54.4	344	2	Q9RVG6	Q9rv66 deinococcus
143	38	55.9	406	2	Q6V0J0	Q6v0j0 verticilliu	216	37	54.4	353	2	Q6UXH1	Q6uxh1 homo sapien
144	38	55.9	415	2	Q98BM3	Q98bm3 rhizobium l	217	37	54.4	353	2	Q86UC0	Q86uc0 homo sapien
145	38	55.9	434	2	Q6FVL2	Q6fvl2 homo sapien	218	37	54.4	366	2	Q9JSG6	Q9jsg6 clostridium
146	38	55.9	496	2	Q6AV41	Q6av41 oryza sativ	219	37	54.4	368	2	Q8P414	Q8p414 xanthomonas
147	38	55.9	504	2	Q6BXX0	Q6bxx0 debaryomyce	220	37	54.4	374	2	Q87754	Q87754 klebsiella
148	38	55.9	513	2	Q82RR9	Q82rr9 streptomyc	221	37	54.4	375	1	Y110_MVCLE	Q49721 mycobacteri
149	38	55.9	555	2	Q7NKA5	Q7nka5 chromobacte	222	37	54.4	375	2	Q94GX2	Q94gx2 oryza sativ
150	38	55.9	569	2	Q9ZBP0	Q9zbp0 streptomyc	223	37	54.4	377	2	Q8ZPL4	Q8zpl4 salmonella
151	38	55.9	580	2	Q9H0W3	Q9h0w3 homo sapien	224	37	54.4	382	1	OMS2_SALTI	Q56111 salmonella
152	38	55.9	589	2	Q66HG7	Q66hg7 rattus norv	225	37	54.4	387	2	Q87ES1	Q87es1 xylella fas
153	38	55.9	601	2	Q93OC1	Q93oc1 rhizobium m	226	37	54.4	388	1	CUDP_METAN	P29138 metarhizium
154	38	55.9	617	2	Q8X959	Q8x959 escherichia	227	37	54.4	388	2	Q96UF4	Q96uf4 metarhizium
155	38	55.9	619	2	Q81WV3	Q81wv3 homo sapien	228	37	54.4	389	2	Q8DJV0	Q8djv0 synschococc
156	38	55.9	619	2	Q9DBN9	Q9dbn9 mus musculu	229	37	54.4	389	2	Q9EGL7	Q9egl7 xylella fas
157	38	55.9	645	2	Q74Z10	Q74z10 ashbya goss	230	37	54.4	390	2	Q96UF8	Q96uf8 metarhizium
158	38	55.9	831	2	Q9PU49	Q9pu49 gallus gall	231	37	54.4	390	2	Q96UG1	Q96ug1 metarhizium
159	38	55.9	843	2	Q6FW78	Q6fw78 candida gla	232	37	54.4	390	2	Q9UUR5	Q9uur5 metarhizium
160	38	55.9	921	2	Q67T45	Q67t45 symbiobacte	233	37	54.4	407	2	Q8FSQ0	Q8fsq0 corynebacte
161	38	55.9	953	2	Q757Q5	Q757q5 ashbya goss	234	37	54.4	409	2	Q9EX13	Q9ex13 deinococcus
162	38	55.9	1048	2	Q9NA71	Q9na71 caenorhabdi	235	37	54.4	413	2	Q6L742	Q6l742 streptomyc
163	38	55.9	2802	2	Q9DER5	Q9der5 gallus gall	236	37	54.4	418	2	Q67NS0	Q67ns0 symbiobacte
164	37.5	55.1	776	2	Q9RJ01	Q9rj01 streptomyc	237	37	54.4	420	2	Q7T094	Q7t094 arabidopsis
165	37	54.4	106	2	Q7WXH6	Q7wxh6 alcaligenes	238	37	54.4	432	2	Q751K5	Q751k5 ashbya goss
166	37	54.4	130	1	CYC3_DESVM	P00132 desulfovibr	239	37	54.4	436	2	Q86484	Q86484 streptomyc
167	37	54.4	145	2	Q73X37	Q73x37 mycobacteri	240	37	54.4	439	2	Q891P7	Q891p7 bradyrhizob
168	37	54.4	167	2	Q9LOW9	Q9low9 streptomyc	241	37	54.4	454	2	Q9SNM0	Q9snm0 oryza sativ
169	37	54.4	173	2	Q8TWZ8	Q8twz8 methanopyru	242	37	54.4	454	2	Q9L4D7	Q9l4d7 xanthomonas
170	37	54.4	186	2	Q8PXD9	Q8pxd9 methanosarc	243	37	54.4	454	2	Q8P9K9	Q8p9k9 xanthomonas
171	37	54.4	198	2	Q89HG8	Q89hg8 bradyrhizob	244	37	54.4	471	2	Q9RC6K	Q9rc6k streptomyc
172	37	54.4	199	2	Q9RXG9	Q9rxg9 deinococcus	245	37	54.4	473	2	Q71211	Q71211 streptomyc
173	37	54.4	211	2	Q7Q127	Q7q127 anopheles g	246	37	54.4	499	2	Q6D3A1	Q6d3a1 erwinia car
174	37	54.4	219	2	Q8KNG4	Q8kng4 micromonospe	247	37	54.4	499	2	Q7VWD9	Q7vwd9 bordetella
175	37	54.4	220	2	Q6IK12	Q6ik12 drosophila	248	37	54.4	499	2	Q7W9N6	Q7w9n6 bordetella
176	37	54.4	224	2	Q8WB45	Q8wb45 todus angus	249	37	54.4	499	2	Q7WH21	Q7wh21 bordetella
177	37	54.4	224	2	Q8WC17	Q8wc17 todus angus	250	37	54.4	506	1	LNT_VIBVU	Q8dfe4 vibrio vuln



251	37	54.4	506	1	LNT_VIBVY	Q7mm04 vibrio vuln	324	36	52.9	206	2	Q9RWM2	Q9rwm2 deinococcus
252	37	54.4	514	2	Q7S0Y2	Q7s0y2 neurospora	325	36	52.9	207	2	Q82H72	Q82h72 streptomyces
253	37	54.4	520	2	Q6NH48	Q6nh48 corynebacte	326	36	52.9	208	2	Q8DRC9	Q8drc9 streptococ
254	37	54.4	528	2	Q968M1	Q968m1 trypanosoma	327	36	52.9	214	2	Q985E9	Q985e9 rhizobium 1
255	37	54.4	539	2	Q91ZX0	Q91zx0 arabidopsis	328	36	52.9	219	2	Q7QDF7	Q7qdf7 anopheles 9
256	37	54.4	551	2	Q9KX16	Q9kx16 streptomyces	329	36	52.9	219	2	Q82L94	Q82l94 streptomyces
257	37	54.4	556	2	Q7U617	Q7u617 synechococ	330	36	52.9	221	2	Q7VZD9	Q7vzd9 bordetella
258	37	54.4	567	1	CC45_XENLA	Q9yhz6 xenopus lae	331	36	52.9	235	2	Q9XTH1	Q9xth1 caenorhabdi
259	37	54.4	567	2	Q6AZM5	Q6azm5 xenopus lae	332	36	52.9	235	2	Q8PCA8	Q8pc8 xanthomonas
260	37	54.4	567	2	Q6GN24	Q6gn24 xenopus lae	333	36	52.9	235	2	Q8PNZ0	Q8pnz0 xanthomonas
261	37	54.4	579	2	Q9REN3	Q9ren3 zymomonas m	334	36	52.9	240	2	Q81857	Q81857 caenorhabdi
262	37	54.4	613	2	Q9FTG6	Q9ftg6 oryza sativ	335	36	52.9	241	2	Q7WGW9	Q7wgv9 bordetella
263	37	54.4	630	2	Q8MTA4	Q8mta4 drosophila	336	36	52.9	243	2	Q8EV53	Q8ev53 mycoplasma
264	37	54.4	637	2	Q6NFB8	Q6nfb8 corynebacte	337	36	52.9	247	2	Q9VZF9	Q9vzf9 drosophila
265	37	54.4	640	2	Q6ZV16	Q6zv16 homo sapien	338	36	52.9	247	2	Q923P5	Q923p5 mus musculu
266	37	54.4	658	2	Q925X8	Q925x8 frankia sp.	339	36	52.9	248	2	Q8D459	Q8d459 vibrio vuln
267	37	54.4	686	2	Q7KTV5	Q7ktv5 drosophila	340	36	52.9	251	2	Q84H76	Q84h76 rhizobium 1
268	37	54.4	699	2	Q8TWT4	Q8tw4 methanopyru	341	36	52.9	251	2	Q9D9K5	Q9d9k5 m mus muscu
269	37	54.4	700	2	Q8N2X7	Q8n2x7 homo sapien	342	36	52.9	252	2	Q9BQ73	Q9bq73 oryza sativ
270	37	54.4	701	1	EFG_MYCTU	O53790 mycobacteri	343	36	52.9	253	2	Q6F3B3	Q6f3b3 oryza sativ
271	37	54.4	715	2	Q7V4P6	Q7v4p6 prochloroco	344	36	52.9	254	2	Q8U9H7	Q8u9h7 agrobacteri
272	37	54.4	720	2	Q7Z2C9	Q7z2c9 drosophila	345	36	52.9	256	2	Q8A2S0	Q8a2s0 caulobacter
273	37	54.4	720	2	Q9VNX5	Q9vnx6 drosophila	346	36	52.9	258	2	Q7W9T8	Q7w9t8 bordetella
274	37	54.4	730	2	Q8BTY5	Q8bt5 homo sapien	347	36	52.9	271	2	Q631Q7	Q631q7 burkholderi
275	37	54.4	788	2	Q97SG8	Q97sg8 streptococ	348	36	52.9	275	2	Q8YYA4	Q8yya4 anabaena sp
276	37	54.4	788	2	Q8DR28	Q8dr28 streptococ	349	36	52.9	284	2	Q8EV55	Q8ev55 mycoplasma
277	37	54.4	800	2	Q8GF31	Q8gf31 zymomonas m	350	36	52.9	286	2	Q95PX0	Q95px0 caenorhabdi
278	37	54.4	802	2	Q8L4U4	Q8l4u4 oryza sativ	351	36	52.9	288	2	Q9VNM7	Q9vnm7 drosophila
279	37	54.4	817	2	Q8NZ13	Q8nzi3 streptococ	352	36	52.9	288	2	Q88149	Q88149 pseudomonas
280	37	54.4	817	2	Q9YF68	Q9yf68 streptococ	353	36	52.9	290	2	Q9F373	Q9f373 streptomyces
281	37	54.4	817	2	Q8K5Y4	Q8k5y4 streptococ	354	36	52.9	291	2	Q6ANW5	Q6anw5 desulfotale
282	37	54.4	860	2	Q8R3N1	Q8r3n1 mus musculu	355	36	52.9	292	2	Q882G7	Q882g7 pseudomonas
283	37	54.4	922	2	Q76384	Q76384 caenorhabdi	356	36	52.9	294	2	Q67L16	Q67l16 symbiobacte
284	37	54.4	922	2	Q8L4V9	Q8l4v9 streptomyces	357	36	52.9	296	1	N111_ARATH	Q8vzw1 arabidopsis
285	37	54.4	952	2	Q988M2	Q988m2 rhizobium 1	358	36	52.9	302	2	Q9N394	Q9n394 caenorhabdi
286	37	54.4	965	2	Q629W8	Q629w8 burkholderi	359	36	52.9	304	2	Q96182	Q96182 homo sapien
287	37	54.4	969	1	FTSK_BIFLO	Q8g4h3 bifidobacte	360	36	52.9	304	2	Q6ZMB1	Q6zmb1 homo sapien
288	37	54.4	988	2	Q631I0	Q631i0 burkholderi	361	36	52.9	306	2	Q8IUL7	Q8iul7 homo sapien
289	37	54.4	1042	2	Q7VSF2	Q7vsf2 bordetella	362	36	52.9	312	2	Q7MFN5	Q7mfn5 vibrio vuln
290	37	54.4	1302	2	Q7UUN6	Q7uun6 rhodopirell	363	36	52.9	315	2	Q8PBA0	Q8pba0 xanthomonas
291	37	54.4	1307	1	BC4_ACEXY	Q9wx71 acetobacter	364	36	52.9	320	2	Q8RMK7	Q8rmk7 azospirillu
292	37	54.4	1312	2	Q6INC1	Q6incl xenopus lae	365	36	52.9	328	2	Q9G8V3	Q9g8v3 rhodomonas
293	37	54.4	1318	2	Q6C425	Q6c425 yarrowia li	366	36	52.9	329	1	PSBO_FRIAG	Q49079 fritillaria
294	37	54.4	1460	2	Q8K6B1	Q8k6b1 streptococ	367	36	52.9	329	2	Q9V2X6	Q9v2x6 methanobact
295	37	54.4	2308	2	Q9VP17	Q9vp17 drosophila	368	36	52.9	331	2	Q8TVS6	Q8tvse methanopyru
296	36.5	53.7	353	2	Q9AG17	Q9ag17 pisum sativ	369	36	52.9	342	2	Q56284	Q56284 treponema d
297	36	52.9	58	2	Q8MLZ0	Q8mlz0 drosophila	370	36	52.9	346	2	Q73KN6	Q73kn6 treponema d
298	36	52.9	58	2	Q8MTG7	Q8mtg7 drosophila	371	36	52.9	348	2	Q8DHJ2	Q8dhj2 synechococ
299	36	52.9	62	2	Q7AKJ8	Q7akj8 streptomyces	372	36	52.9	358	1	CHAD_MOUSE	O55226 mus musculu
300	36	52.9	62	2	Q9KIM3	Q9kim3 streptomyces	373	36	52.9	358	1	CHAD_RAT	O70210 rattus norv
301	36	52.9	70	2	Q8EGV4	Q8egv4 shewanella	374	36	52.9	359	2	Q9F2J7	Q9f2j7 streptomyces
302	36	52.9	84	2	Q8Z6Q1	Q8z6q1 salmonella	375	36	52.9	369	2	Q82NU9	Q82nu9 streptomyces
303	36	52.9	93	2	Q8P3P8	Q8p3p8 xanthomonas	376	36	52.9	370	2	Q88JL2	Q88jl2 pseudomonas
304	36	52.9	109	1	YDHA_ECOLI	P28224 escherichia	377	36	52.9	380	2	Q9KYD1	Q9kyd1 streptomyces
305	36	52.9	111	2	Q7QLD7	Q7ql7 anopheles g	378	36	52.9	387	1	TRAB_RHISN	P55416 rhizobium s
306	36	52.9	123	2	Q6F3B2	Q6f3b2 oryza sativ	379	36	52.9	389	2	O85150	O85150 ralatonia s
307	36	52.9	140	2	Q6EPQ7	Q6epq7 oryza sativ	380	36	52.9	389	2	O8PFJ6	O8pfj6 xanthomonas
308	36	52.9	151	2	Q8ANS4	Q8ans4 rhodobacter	381	36	52.9	395	2	Q95PW9	Q95pw9 caenorhabdi
309	36	52.9	153	2	Q9VLB1	Q9vlb1 drosophila	382	36	52.9	395	2	O6NDB3	O6nd83 rhodopseudo
310	36	52.9	155	2	Q7MK08	Q7mk08 vibrio vuln	383	36	52.9	402	2	Q88F24	Q88f24 pseudomonas
311	36	52.9	155	2	Q8DA71	Q8da71 vibrio vuln	384	36	52.9	403	2	O6N454	O6n454 rhodopseudo
312	36	52.9	173	2	Q6CSK2	Q6csk2 kluyveromyc	385	36	52.9	408	2	O8KGV3	O8kgv3 rhizobium 1
313	36	52.9	183	2	O54022	O54022 paracoccus	386	36	52.9	409	2	O8NXS1	O8nxs1 staphylococ
314	36	52.9	185	2	Q89H01	Q89h01 bradyrhizob	387	36	52.9	409	2	Q7A2V6	Q7a2v6 staphylococ
315	36	52.9	192	2	Q8TTR2	Q8tr2 methanosarc	388	36	52.9	409	2	Q7A706	Q7a706 staphylococ
316	36	52.9	195	2	Q8NZ09	Q8nz09 streptococ	389	36	52.9	409	2	O9X668	O9x668 staphylococ
317	36	52.9	195	2	Q99XJ8	Q99xj8 streptococ	390	36	52.9	409	2	O6GBI5	O6gb15 staphylococ
318	36	52.9	195	2	O8K5I1	O8k5i1 streptococ	391	36	52.9	409	2	O6GJ25	O6gj25 staphylococ
319	36	52.9	198	2	O652K5	Q652k5 oryza sativ	392	36	52.9	411	2	O53873	O53873 streptomyces
320	36	52.9	202	1	O68G_RABIT	Q28679 oryctolagus	393	36	52.9	411	2	Q9RU03	Q9ru03 deinococcus
321	36	52.9	202	2	Q8THA7	Q8tha7 methanosarc	394	36	52.9	417	2	Q7NSV5	Q7nsv5 chromobacte
322	36	52.9	202	2	Q989Y6	Q989y6 rhizobium 1	395	36	52.9	419	1	Y4ML_RHISN	P55571 rhizobium s
323	36	52.9	204	2	Q67L32	Q67l32 symbiobacte	396	36	52.9	419	2	Q8S1V0	Q8s1v0 oryza sativ

397	36	52.9	428	2	Q9ABC9	Q9abc9 caulobacter	470	35	51.5	130	1	LY6E_MOUSE	Q64253 mus musculus
398	36	52.9	481	2	Q8C6T5	Q8c6t5 mus musculus	471	35	51.5	130	2	Q6T6F7	Q6t6f7 brachydanio
399	36	52.9	486	2	Q94HE8	Q94he8 oryza sativ	472	35	51.5	133	2	Q9CKN2	Q9ckn2 mus musculus
400	36	52.9	486	2	Q7XF13	Q7xf13 oryza sativ	473	35	51.5	136	2	Q6AY73	Q6ay73 rattus norv
401	36	52.9	486	2	Q6MPG8	Q6mpg8 bdellovibri	474	35	51.5	136	2	Q99JA5	Q99ja5 mus musculus
402	36	52.9	488	2	Q922C8	Q922c8 rhizobium m	475	35	51.5	138	2	Q9YFN2	Q9yfn2 aeropyrum p
403	36	52.9	492	2	Q6N3X0	Q6n3x0 rhodopseudo	476	35	51.5	141	2	Q8PGQ8	Q8pgq8 xanthomonas
404	36	52.9	501	2	Q8XX19	Q8xx19 ralstonia s	477	35	51.5	143	2	Q8FJ84	Q8fj84 escherichia
405	36	52.9	506	2	Q67MP2	Q67mp2 symbiobacte	478	35	51.5	147	2	Q9KRH9	Q9krh9 vibrio chol
406	36	52.9	516	2	Q7PIL5	Q7pil5 chromobacte	479	35	51.5	148	1	AZUR_PSEAE	P00282 pseudomonas
407	36	52.9	524	2	Q643R3	Q643r3 homo sapien	480	35	51.5	150	2	Q7U775	Q7u775 synechococ
408	36	52.9	576	1	PYRG_ASHGO	Pyrg5117 ashbya goss	481	35	51.5	155	2	Q6NI52	Q6ni52 corynebacte
409	36	52.9	576	2	Q6DCW8	Q6dcw8 xenopus lae	482	35	51.5	159	1	CU57_ARADI	P80519 araneus dia
410	36	52.9	582	2	Q7NH33	Q7nh33 gloebacter	483	35	51.5	161	2	Q6SGJ8	Q6sgj8 bacillus li
411	36	52.9	584	2	Q8PE94	Q8pe94 xanthomonas	484	35	51.5	161	2	Q91053	Q91053 pseudomonas
412	36	52.9	587	2	Q89DA2	Q89da2 bradyrhizob	485	35	51.5	171	2	Q6AQ77	Q6aq77 desulfotale
413	36	52.9	598	2	Q8AA03	Q8aa03 bacteroides	486	35	51.5	176	2	Q8UG90	Q8ug90 agrobacteri
414	36	52.9	603	2	Q7M994	Q7m994 wolinnella s	487	35	51.5	188	2	Q62B89	Q62b89 burkholderi
415	36	52.9	615	2	Q8YRA5	Q8yra5 anabaena sp	488	35	51.5	191	2	Q97064	Q97064 drosophila
416	36	52.9	623	2	Q9XA73	Q9xa73 streptomyce	489	35	51.5	191	2	Q7ME86	Q7me86 vibrio vuln
417	36	52.9	625	2	Q9VIL5	Q9vil5 drosophila	490	35	51.5	191	2	Q8D782	Q8d782 vibrio vuln
418	36	52.9	657	2	Q8S5W4	Q8s5w4 oryza sativ	491	35	51.5	199	2	Q63NI0	Q63ni0 burkholderi
419	36	52.9	658	2	Q8PR52	Q8pr52 xanthomonas	492	35	51.5	202	2	Q8TDY7	Q8tdy7 homo sapien
420	36	52.9	663	2	Q7QI59	Q7qi59 anopheles g	493	35	51.5	202	2	Q9N9A7	Q9n9a7 leishmania
421	36	52.9	673	2	Q82BP9	Q82bp9 streptomyce	494	35	51.5	204	2	Q7MTX5	Q7mtx5 porphyromon
422	36	52.9	682	2	Q9A8Q8	Q9a8q8 caulobacter	495	35	51.5	205	2	Q70WV0	Q70wv0 spintharina
423	36	52.9	722	2	Q9VWF8	Q9vwp8 drosophila	496	35	51.5	206	2	Q63LZ1	Q63lz1 burkholderi
424	36	52.9	732	2	Q9L281	Q9l281 streptomyce	497	35	51.5	206	2	Q8WU78	Q8wuy8 bos taurus
425	36	52.9	769	1	DCB2_MOUSE	Q912v3 mus musculus	498	35	51.5	206	2	Q8NS72	Q8ns72 homo sapien
426	36	52.9	769	1	DCB2_RAT	Q912v2 rattus norv	499	35	51.5	206	2	Q8MRD2	Q8mrdr2 drosophila
427	36	52.9	800	2	Q8ZXD1	Q8zxd1 pyrobaculum	500	35	51.5	206	2	Q9N0D0	Q9n0d0 macaca fasc
428	36	52.9	811	2	Q7UUN7	Q7uun7 rhodopirell	501	35	51.5	206	2	Q62BF0	Q62bf0 burkholderi
429	36	52.9	819	2	Q86GI8	Q86gi8 pseudomonas	502	35	51.5	206	2	Q63LZ1	Q63lz1 burkholderi
430	36	52.9	832	2	Q8WU28	Q8wu28 homo sapien	503	35	51.5	206	2	Q8BVG8	Q8bvvg8 m mus muscu
431	36	52.9	1028	1	CA16_HUMAN	P12109 homo sapien	504	35	51.5	208	2	Q97062	Q97062 drosophila
432	36	52.9	1028	2	Q7Z645	Q7z645 homo sapien	505	35	51.5	210	2	Q7RYL3	Q7ryl3 neuropeptora
433	36	52.9	1096	2	Q6R524	Q6r524 oreochromis	506	35	51.5	210	2	Q6L5A5	Q6l5a5 oryza sativ
434	36	52.9	1165	2	Q7W3X5	Q7w3x5 bordetella	507	35	51.5	211	2	Q9GC11	Q9gc11 wuchereria
435	36	52.9	1168	2	Q7WFA5	Q7wfa5 bordetella	508	35	51.5	211	2	Q9GC47	Q9gc47 buglia mala
436	36	52.9	1325	2	Q819N3	Q819n3 anopheles g	509	35	51.5	214	2	Q7ACY3	Q7acy3 escherichia
437	36	52.9	1706	2	Q42900	Q42900 schizosacch	510	35	51.5	224	2	Q69836	Q69836 streptomyce
438	36	52.9	1783	1	RAA3_CHLRE	Q9fec4 chlamydomon	511	35	51.5	224	2	Q8XB87	Q8xb87 escherichia
439	36	52.9	1999	2	Q80U36	Q80u36 mus musculus	512	35	51.5	227	2	Q7J3Z2	Q7j3z2 scomberomor
440	36	52.9	2666	2	Q6FW99	Q6fw99 candida gla	513	35	51.5	227	2	Q7J3Z8	Q7j3z8 scomberomor
441	36	52.9	2666	2	Q7PFP1	Q7ppf1 anopheles g	514	35	51.5	227	2	Q7J405	Q7j405 scomberomor
442	36	52.9	4594	1	DYHC_HUMAN	Q14204 homo sapien	515	35	51.5	227	2	Q7U409	Q7u409 scomberomor
443	36	52.9	4644	1	DYHC_MOUSE	Q9jhu4 mus musculus	516	35	51.5	227	2	Q9TGC4	Q9tgc4 scomberomor
444	36	52.9	4644	1	DYHC_RAT	P38650 rattus norv	517	35	51.5	227	2	Q9TGC5	Q9tgc5 scomberomor
445	36	52.9	4646	2	Q6DKQ7	Q6dkq7 homo sapien	518	35	51.5	227	2	Q9TGC6	Q9tgc6 scomberomor
446	35.5	52.2	289	2	Q9YDR1	Q9ydr1 aeropyrum p	519	35	51.5	227	2	Q9TGC7	Q9tgc7 scomberomor
447	35.5	52.2	297	2	Q71BE9	Q71be9 barley yell	520	35	51.5	227	2	Q9TGC8	Q9tgc8 scomberomor
448	35.5	52.2	306	2	Q8EHT8	Q8eht8 shewanella	521	35	51.5	227	2	Q9XJZ4	Q9xjz4 scomberomor
449	35.5	52.2	2412	1	POL1_BAYMG	Q04574 b genome po	522	35	51.5	227	2	Q9XJZ6	Q9xjz6 scomberomor
450	35	51.5	64	2	Q7QQ78	Q7qqt8 giardia lam	523	35	51.5	227	2	Q9XKZ7	Q9xkz7 scomberomor
451	35	51.5	83	2	Q6H7B7	Q6h7b7 oryza sativ	524	35	51.5	227	2	Q9XPA4	Q9xpa4 scomberomor
452	35	51.5	83	2	Q47857	Q47857 pantoea agg	525	35	51.5	227	2	Q9XPA9	Q9xpa9 scomberomor
453	35	51.5	86	2	Q46647	Q46647 erwinia amy	526	35	51.5	227	2	Q9XPB0	Q9xpb0 scomberomor
454	35	51.5	88	1	CU70_LOCMI	P80232 locusta mig	527	35	51.5	227	2	Q9XPB1	Q9xpb1 scomberomor
455	35	51.5	94	2	Q9XUV3	Q9xuv3 caenorhabdi	528	35	51.5	228	2	Q82AB8	Q82ab8 streptomyce
456	35	51.5	94	2	Q89I78	Q89i78 bradyrhizob	529	35	51.5	231	2	Q9PZ39	Q9pz39 xestia c-ni
457	35	51.5	94	2	Q8JZT1	Q8jzt1 mus musculus	530	35	51.5	232	2	Q7NQB4	Q7nqb4 chromobacte
458	35	51.5	107	2	Q89KQ4	Q89kq4 bradyrhizob	531	35	51.5	234	2	Q9I656	Q9i656 pseudomonas
459	35	51.5	109	2	Q82FNO	Q8zpn0 salmonella	532	35	51.5	238	2	Q666T0	Q666t0 yersinia ps
460	35	51.5	113	2	Q9GUZ4	Q9guz4 aurelia aur	533	35	51.5	238	2	Q8ZHK2	Q8zhk2 yersinia pe
461	35	51.5	114	2	Q9GUY9	Q9guy9 tima formos	534	35	51.5	239	1	YC43_CVACA	Q2185 cyanidium c
462	35	51.5	114	2	Q9GUZ0	Q9guz0 hydractinia	535	35	51.5	244	2	Q6A5C0	Q6a5c0 propionibac
463	35	51.5	114	2	Q9GUZ2	Q9guz2 craepedacus	536	35	51.5	248	2	Q6NJS9	Q6nj9 corynebacte
464	35	51.5	115	1	HPY2_AGABI	P49073 agarius bi	537	35	51.5	254	2	Q6N818	Q6n818 rhodopseudo
465	35	51.5	115	2	Q9ATH2	Q9ath2 corylus ave	538	35	51.5	255	2	Q8ZL13	Q8z13 pyrobaculum
466	35	51.5	117	2	Q7XZE0	Q7xze0 atriplex nu	539	35	51.5	255	2	Q7NUS5	Q7nus5 chromobacte
467	35	51.5	119	2	Q6L4S7	Q6l4s7 oryza sativ	540	35	51.5	255	2	Q76C28	Q76c28 fugu rubrip
468	35	51.5	126	2	Q9I2H0	Q9i2h0 pseudomonas	541	35	51.5	261	2	Q7NKI5	Q7nk15 gloeobacter
469	35	51.5	128	2	Q6V9R2	Q6v9r2 penaeus mon	542	35	51.5	262	2	Q93PZ0	Q93pz0 pseudomonas

543	35	51.5	263	1	YC43_ODOSI	P49538 odontella s	616	35	51.5	424	2	Q9KZ99	Q9KZ99 streptomyc
544	35	51.5	267	2	Q7NKT1	Q7nkt1 gloebacter	617	35	51.5	426	2	Q7ZZW1	Q7zzw1 oncorhynch
545	35	51.5	267	2	Q8UEQ1	Q8ueq1 agrobacteri	618	35	51.5	429	2	Q6MXM1	Q6mxm1 serratia ma
546	35	51.5	271	2	Q8DUY4	Q8duy4 streptococc	619	35	51.5	429	2	Q9L336	Q9l336 serratia ma
547	35	51.5	272	2	Q72DG3	Q72dg3 desulfovibr	620	35	51.5	441	2	Q6D210	Q6d210 erwinia car
548	35	51.5	274	2	Q67LRS	Q67lrs symbiobacte	621	35	51.5	451	2	Q9CK72	Q9ck72 pasteurella
549	35	51.5	274	2	Q6AIR5	Q6air5 desulfonema	622	35	51.5	453	1	ANML_HUMAN	Q6bxtj7 homo sapien
550	35	51.5	276	2	Q6J501	Q6j501 steinernema	623	35	51.5	453	2	Q6LNT7	Q6lnt7 photobacter
551	35	51.5	280	1	BTUP_YERPE	Q8zbm3 yersinia pe	624	35	51.5	463	2	Q9EYGO	Q9eygo amycolatops
552	35	51.5	284	2	Q67RJ7	Q67rj7 symbiobacte	625	35	51.5	464	1	NORM_PASNU	Q9cmz9 pasteurella
553	35	51.5	285	2	Q8FL09	Q8fl09 escherichia	626	35	51.5	468	2	Q82GV7	Q82gv7 streptomyc
554	35	51.5	286	2	Q6ARK3	Q6ark3 desulfotale	627	35	51.5	472	2	Q8S5P9	Q8s5p9 oryza sativ
555	35	51.5	287	2	Q88DF0	Q88df0 pseudomonas	628	35	51.5	480	2	Q8RUV5	Q8ruv5 oryza sativ
556	35	51.5	291	2	Q82E64	Q82e64 streptomyc	629	35	51.5	482	2	Q7UPY6	Q7upy6 rhodopirell
557	35	51.5	291	2	Q8E478	Q8e478 streptococc	630	35	51.5	492	2	Q6JAD9	Q6jad9 zea mays m
558	35	51.5	292	2	Q74H73	Q74h73 geobacter s	631	35	51.5	496	2	Q7WNM4	Q7wn4 bordetella
559	35	51.5	295	2	Q8DYM1	Q8dym1 streptococc	632	35	51.5	496	2	Q7WH84	Q7wh84 bordetella
560	35	51.5	296	2	Q89048	Q89048 trypanosoma	633	35	51.5	498	2	Q7WFB2	Q7wfb2 bordetella
561	35	51.5	297	2	Q826C1	Q826c1 streptomyc	634	35	51.5	500	2	Q9YDJ6	Q9ydj6 aeropyrum p
562	35	51.5	302	2	Q8UHR8	Q8uhr8 agrobacteri	635	35	51.5	507	2	Q8M0D0	Q8m0d0 amoebidium
563	35	51.5	307	2	Q746U8	Q746u8 geobacter s	636	35	51.5	507	2	Q810L1	Q810l1 mus musculu
564	35	51.5	307	2	Q7MKK1	Q7mkk1 vibrio vuln	637	35	51.5	509	2	Q7NI25	Q7ni25 gloebacter
565	35	51.5	310	1	DHYS_THEAC	Q9hl74 thermoplaem	638	35	51.5	510	2	Q8NTN9	Q8ntn9 corynebacte
566	35	51.5	310	2	Q9XX39	Q9xx39 caenorhabdi	639	35	51.5	518	2	Q9ZVS8	Q9zvs8 arabidopsis
567	35	51.5	312	2	Q9RIV5	Q9riv5 streptomyc	640	35	51.5	518	2	Q74B39	Q74b39 geobacter s
568	35	51.5	312	2	Q7NTA5	Q7nta5 chromobacte	641	35	51.5	519	1	PPBT_CHICK	Q92058 gallus gall
569	35	51.5	315	2	Q6H7C9	Q6h7c9 oryza sativ	642	35	51.5	520	2	Q92Q56	Q92qe6 rhizobium m
570	35	51.5	320	2	Q99047	Q99047 trypanosoma	643	35	51.5	524	2	Q66JPT	Q66jpt mus musculu
571	35	51.5	321	2	Q7D138	Q7d138 agrobacteri	644	35	51.5	524	2	Q6NVGL	Q6nvg1 mus musculu
572	35	51.5	325	2	Q89YK8	Q89yk8 bacteroides	645	35	51.5	525	2	Q88M35	Q88m35 pseudomonas
573	35	51.5	327	2	Q6IND9	Q6ind9 xenopus lae	646	35	51.5	533	2	Q821R9	Q821r9 salmonella
574	35	51.5	328	2	Q9XA64	Q9xa64 streptomyc	647	35	51.5	535	2	O52716	O52716 klebsiella
575	35	51.5	330	2	Q842G5	Q842g5 pseudomonas	648	35	51.5	549	2	Q9HY61	Q9hy61 pseudomonas
576	35	51.5	334	2	Q7BN99	Q7bn99 pseudomonas	649	35	51.5	552	2	Q66K15	Q66ki5 xenopus lae
577	35	51.5	334	2	Q88LX3	Q88lx3 pseudomonas	650	35	51.5	554	2	Q87A60	Q87a60 xyliella fas
578	35	51.5	335	2	Q8MXL9	Q8mxl9 dictyosteli	651	35	51.5	554	2	Q9PAC4	Q9pac4 xyliella fas
579	35	51.5	341	2	Q7N4Z0	Q7n4z0 photorhabdu	652	35	51.5	557	2	Q82LE0	Q82le0 streptomyc
580	35	51.5	345	2	Q8SJA7	Q8sja7 dodecaceria	653	35	51.5	566	1	CC45_HUMAN	CC4519 homo sapien
581	35	51.5	350	1	ZPB1_MOUSE	Q62522 mus musculu	654	35	51.5	566	1	CC45_MOUSE	CC4519 mus musculu
582	35	51.5	351	2	Q93JK6	Q93jk6 streptomyc	655	35	51.5	566	2	Q6UW54	Q6uw54 homo sapien
583	35	51.5	357	1	PUR5_RHILE	Q9xat2 rhizobium l	656	35	51.5	566	2	Q9UP68	Q9up68 homo sapien
584	35	51.5	359	1	CR12_MOUSE	Q8c119 mus musculu	657	35	51.5	573	2	Q8W0Q9	Q8w0q9 sorghum bic
585	35	51.5	360	2	Q9F889	Q9f889 salmonella	658	35	51.5	584	2	Q76L19	Q76l19 coturnix co
586	35	51.5	363	1	OMPF_SALTI	Q56113 salmonella	659	35	51.5	586	2	Q8H094	Q8h094 oryza sativ
587	35	51.5	363	1	OMPF_SALTU	P37432 salmonella	660	35	51.5	589	2	Q9M4V5	Q9m4v5 oryza sativ
588	35	51.5	363	2	Q8K3E7	Q8k3e7 salmonella	661	35	51.5	591	2	Q7SGN9	Q7sgn9 neurospora
589	35	51.5	368	2	Q8MPT0	Q8mpt0 caenorhabdi	662	35	51.5	591	2	Q65SY7	Q65sy7 manheimia
590	35	51.5	368	2	Q740P8	Q740p8 mycobacteri	663	35	51.5	595	2	Q9C2K3	Q9c2k3 neurospora
591	35	51.5	369	2	Q12593	Q12593 cryptonectr	664	35	51.5	598	2	Q63KC4	Q63kc4 burkholderi
592	35	51.5	373	2	Q6CYA0	Q6cya0 kluyveromyc	665	35	51.5	599	2	Q67RH1	Q67rh1 symbiobacte
593	35	51.5	375	2	Q8CVY6	Q8cvy6 escherichia	666	35	51.5	604	2	Q7S783	Q7s783 neurospora
594	35	51.5	376	2	Q94301	Q94301 caenorhabdi	667	35	51.5	606	2	Q6W282	Q6w2g2 rhizobium s
595	35	51.5	376	2	Q34759	Q54759 synecococc	668	35	51.5	638	2	Q52582	Q52582 ralstonia s
596	35	51.5	382	2	Q68GV9	Q68gv9 lecanicilli	669	35	51.5	638	2	Q8XWL8	Q8xwl8 ralstonia s
597	35	51.5	387	2	Q8TFB2	Q8tfb2 cordyceps c	670	35	51.5	643	2	Q9LR59	Q9lr59 arabidopsis
598	35	51.5	388	2	Q7N624	Q7n624 photorhabdu	671	35	51.5	652	2	Q93H25	Q93h25 streptomyc
599	35	51.5	394	1	OMS1_SALTI	Q56110 salmonella	672	35	51.5	659	2	Q96069	Q96069 ciona intes
600	35	51.5	395	2	Q7UE35	Q7ue35 rhodopirell	673	35	51.5	669	2	Q96068	Q96068 ciona intes
601	35	51.5	396	2	Q870Y6	Q870y6 neurospora	674	35	51.5	670	2	Q72GR2	Q72gr2 thermus the
602	35	51.5	397	2	Q97TN9	Q97tn9 clostridium	675	35	51.5	675	2	Q9Y110	Q9y110 desophilla
603	35	51.5	397	1	YEDS_ECOLI	P76335 escherichia	676	35	51.5	677	2	Q51544	Q51544 pseudomonas
604	35	51.5	398	2	Q8ZNS7	Q8zns7 salmonella	677	35	51.5	682	2	Q6Z5K9	Q6z5k9 oryza sativ
605	35	51.5	401	2	Q9KEG8	Q9keg8 bacillus ha	678	35	51.5	684	2	Q986M4	Q986m4 rhizobium l
606	35	51.5	403	2	Q65ED5	Q65ed5 bacillus li	679	35	51.5	688	2	Q9GNP2	Q9gnp2 ciona savig
607	35	51.5	403	2	Q9RSZ7	Q9rsz7 bradyrhizob	680	35	51.5	688	2	Q42375	Q42375 brachydanio
608	35	51.5	407	2	Q9R8R0	Q9r8r0 pseudomonas	681	35	51.5	688	2	Q6PBB2	Q6pbb2 brachydanio
609	35	51.5	407	2	Q88F80	Q88f80 pseudomonas	682	35	51.5	690	1	VG42_BPMU	Q9ct1v6 bacterioph
610	35	51.5	409	1	ALP_TRIHA	Q03420 trichoderma	683	35	51.5	695	2	Q9N602	Q9n602 desophilla
611	35	51.5	409	2	Q86ZV3	Q86zv3 trichoderma	684	35	51.5	699	2	Q9N6P1	Q9n6p1 oryza sativ
612	35	51.5	409	2	Q874K4	Q874k4 trichoderma	685	35	51.5	701	2	Q7UQ82	Q7uq82 rhodopirell
613	35	51.5	417	2	Q7M9Q8	Q7m9q8 wolinnella s	686	35	51.5	702	2	Q84S61	Q84s61 oryza sativ
614	35	51.5	418	2	Q9A4Y7	Q9a4y7 caulobacter	687	35	51.5	713	2	Q6GL27	Q6gl27 xenopus tro
615	35	51.5	423	2	Q8YSL2	Q8ysl2 anabaena sp	688	35	51.5	728	2	Q8MU80	Q8mu80 mastotermes

```

689 35 51.5 738 2 Q6DFA9
690 35 51.5 741 2 O85873
691 35 51.5 758 2 O6TECO
692 35 51.5 770 2 Q9GNP1
693 35 51.5 791 2 Q87V39
694 35 51.5 798 2 Q751L7
695 35 51.5 817 2 Q7QIN8
696 35 51.5 821 2 O8TVF3
697 35 51.5 825 2 Q9RWY0
698 35 51.5 836 2 Q80UG9
699 35 51.5 842 1 VGLH BHV1C
700 35 51.5 848 2 Q6X233
701 35 51.5 849 2 Q9WQY5
702 35 51.5 863 1 IMB1 SCHPO
703 35 51.5 865 2 Q83OV6
704 35 51.5 870 2 O8MRK9
705 35 51.5 913 2 Q6PF28
706 35 51.5 915 2 Q8ABD0
707 35 51.5 916 2 Q6ARD7
708 35 51.5 918 2 Q9BQW5
709 35 51.5 950 2 Q88FQ5
710 35 51.5 956 2 Q69N36
711 35 51.5 970 2 Q8P9U8
712 35 51.5 980 2 Q9ZUI0
713 35 51.5 985 1 CAPP RALSO
714 35 51.5 1018 2 Q8FMEO
715 35 51.5 1075 2 Q73RD3
716 35 51.5 1099 2 Q7MT47
717 35 51.5 1105 2 Q8NIA3
718 35 51.5 1189 2 O85152
719 35 51.5 1235 2 Q6PGG7
720 35 51.5 1290 2 Q82L13
721 35 51.5 1307 2 Q9VW60
722 35 51.5 1314 2 Q8C0S4
723 35 51.5 1319 2 Q9D5Q9
724 35 51.5 1325 1 SOS1 MOUSE
725 35 51.5 1332 2 O8BP86
726 35 51.5 1333 1 SOS1 HUMAN
727 35 51.5 1381 1 VGLP_BEV
728 35 51.5 2055 2 Q9WTS5
729 35 51.5 2764 2 Q9RIK2
730 35 51.5 2765 2 Q63N87
731 35 51.5 4131 2 Q80UC6
732 34.5 50.7 104 2 Q922S9
733 34.5 50.7 142 2 Q6LSD5
734 34.5 50.7 151 2 Q686C4
735 34.5 50.7 151 2 Q686F1
736 34.5 50.7 151 2 Q70772
737 34.5 50.7 169 2 Q6K420
738 34.5 50.7 291 2 FMT CAUCR
739 34.5 50.7 308 2 Q8ALQ2
740 34.5 50.7 365 2 Q622 HUMAN
741 34.5 50.7 368 1 Q8TAM0
742 34.5 50.7 368 2 Q7NI59
743 34.5 50.7 384 2 Q988U4
744 34.5 50.7 617 2 Q8YOC4
745 34.5 50.7 847 2 Q8YOC4
746 34 50.0 18 2 Q9QWG4
747 34 50.0 34 2 Q6B935
748 34 50.0 37 1 PSBY CYACA
749 34 50.0 37 1 PSBY GUITH
750 34 50.0 53 2 Q6IM18

```

## ALIGNMENTS

```

RESULT 1
Q6ND81 PRELIMINARY; PRT; 133 AA.
ID Q6ND81;
AC Q6ND81;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)

```

```

DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein precursor.
GN OrderedLocusNames=RPA0228;
OS Rhodopseudomonas palustris.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Rhodopseudomonas.
OX NCBI_TaxID=1076;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CGA009 / ATCC BAA-98;
RX PubMed=14704707; DOI=10.1038/nbt923;
RA Larimer F.W., Chain P., Hauser L., Lamerdin J.E., Malfatti S., Do L.,
RA Land M.L., Pelletier D.A., Beatty J.T., Lang A.S., Tabita F.R.,
RA Gibson J.L., Hanson T.E., Bobst C., Torres Y. Torres J.L., Perea C.,
RA Harrison F.H., Gibson J., Harwood C.S.;
RT "Complete genome sequence of the metabolically versatile
RT photosynthetic bacterium Rhodopseudomonas palustris.";
RL Nat. Biotechnol. 22:55-61(2004).
DR EMBL; EX572593; CAE25672.1; -.
KW Complete proteome; Hypothetical protein; Signal.
FT SIGNAL 1 37
SQ SEQUENCE 133 AA; 14481 MW; E22F622199718B5B CRC64;

Query Match 66.2%; Score 45; DB 2; Length 133;
Best Local Similarity 78.6%; Pred. No. 15;
Matches 11; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 KKAAYVLLPVLAA 14
Db 12 KSTAARLLPVLAA 25

RESULT 2
Q8H8C7 PRELIMINARY; PRT; 356 AA.
ID Q8H8C7;
AC Q8H8C7;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein OJ1006F06.19.
GN Name=OJ1006F06.19;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA Wing R.A., Yu Y., Soderlund C., Kim H.-R., Rambo T., Saski C.,
RA Currie J., Collura K.;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC099399; AAN0509.1; -.
DR Gramene; O8H8C7; -.
DR GO; GO:0016998; P:cell wall catabolism; IEA.
DR InterPro; IPR002482; LysM.
DR Pfam; PF01476; LysM; 2.
DR SMART; SM00257; LysM; 2.
KW Hypothetical protein.
SQ SEQUENCE 356 AA; 37200 MW; 007D7BED10AB98FE CRC64;

Query Match 66.2%; Score 45; DB 2; Length 356;
Best Local Similarity 84.6%; Pred. No. 36;
Matches 11; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 AAAYVLLPVLAAAP 15
Db 13 AAALLLVLLAAAP 25

RESULT 3
Q6N2W2 PRELIMINARY; PRT; 497 AA.
ID Q6N2W2
AC Q6N2W2;

```

```

DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein precursor.
GN OrderedLocusNames=RPA3936;
OS Rhodopseudomonas palustris.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiacae; Rhodopseudomonas.
OX NCBI_TaxID=1076;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CGA009 / ATCC BAA-98;
RX PubMed=14704707; DOI=10.1038/nbt923;
RA Larimer F.W., Chain P., Hauser L., Lamerdin J.E., Malfatti S., Do L.,
RA Land M.L., Pelletier D.A., Beatty J.T., Lang A.S., Tabita F.R.,
RA Gibson J.L., Hanson T.E., Bobst C., Torres y Torres J.L., Peres C.,
RA Harrison F.H., Gibson J., Harwood C.S.;
RT "Complete genome sequence of the metabolically versatile
RT photosynthetic bacterium Rhodopseudomonas palustris.";
RL Nat. Biotechnol. 22:55-61(2004).
DR EMBL; BX572605; CAE29377.1; -.
KW Complete proteome; Hypothetical protein; Signal.
FT SIGNAL 1 30 Potential.
SQ SEQUENCE 497 AA; 54125 MW; 6A764989856F441E CRC64;

Query Match 64.7%; Score 44; DB 2; Length 497;
Best Local Similarity 75.0%; Pred. No. 73;
Matches 9; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 AAALLPVLAA 14
Db 19 AAAILPVLAA 30
|||||
|:::|

RESULT 4
Q87TJ5 ID Q87TJ5 PRELIMINARY; PRT; 190 AA.
AC Q87TJ5;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein VP0074.
GN OrderedLocusNames=VP0074;
OS Vibrio parahaemolyticus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=670;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RIMD 2210633 / Serotype O3:K6;
RX MEDLINE=22508454; PubMed=12620739; DOI=10.1016/S0140-6736(03)12659-1;
RA Makino K., Oshima K., Kurokawa K., Yokoyama K., Uda T., Tagomori K.,
RA Iijima Y., Nejima M., Nakano M., Yamashita A., Kubota Y., Kimura S.,
RA Yasunaga T., Honda T., Shinagawa H., Hattori M., Iida T.;
RT "Genome sequence of Vibrio parahaemolyticus: a pathogenic mechanism
RT distinct from that of V. cholerae.";
RL Lancet 361:743-749(2003).
DR EMBL; AP05073; BAC58337.1; -.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 190 AA; 20907 MW; 07064D073E1FA542 CRC64;

Query Match 63.2%; Score 43; DB 2; Length 190;
Best Local Similarity 60.0%; Pred. No. 44;
Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 KKAALVLPVLAAP 15
Db 103 KQAQKVTVPSNAAP 117
|:::|
|:::|

RESULT 5
Q8FRS4 ID Q8FRS4 PRELIMINARY; PRT; 340 AA.

```

```

AC Q8FRS4;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative iron transport membrane protein.
GN OrderedLocusNames=CR0685;
OS Corynebacterium efficiens.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxID=152794;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=YS-314;
RX MEDLINE=22723752; PubMed=12840036; DOI=10.1101/gr.1285603;
RA Nishio Y., Nakamura Y., Kawarabayashi Y., Ueuda Y., Kimura E.,
RA Sugimoto S., Matsui K., Yamagishi A., Kikuchi H., Ikeo K.,
RA Gojohori T.;
RT "Comparative complete genome sequence analysis of the amino acid
RT replacements responsible for the thermostability of Corynebacterium
RT efficiens.";
RL Genome Res. 13:1572-1579(2003).
DR EMBL; AP005216; BAC17495.1; -.
DR HSRF; P06609; IL7V.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005215; P:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000522; FcCD.
DR Pfam; PF01032; FcCD; 1.
KW Complete proteome.
SQ SEQUENCE 340 AA; 34818 MW; 14B047C6A943A576 CRC64;

Query Match 63.2%; Score 43; DB 2; Length 340;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 AVLLPVLAA 14
Db 211 AVLLPVLAA 220
|||||
|:::|

RESULT 6
Q99PG1 ID Q99PG1 PRELIMINARY; PRT; 582 AA.
AC Q99PG1;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Matrix metalloprotease MT1-MMP.
GN Name=MMP14;
OS Cricetus griseus (Chinese hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Cricetulus.
OX NCBI_TaxID=10029;
RN [1]
RP SEQUENCE FROM N.A.
RA Borrell-Pages M., Arribas J.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF306662; AAG34676.2; -.
DR HSRF; P50281; 1BQQ.
DR MEROPS; M10.014; -.
DR GO; GO:0005578; C:extracellular matrix (sensu Metazoa); IEA.
DR GO; GO:0004222; P:metalloendopeptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR000585; Hemopepin.
DR InterPro; IPR001818; Pept_M10A_M12B.
DR InterPro; IPR006025; Pept_M_Zn_BS.
DR InterPro; IPR009070; PGED_like.
DR Pfam; PF00045; Hemopepin; 4.
DR Pfam; PF03933; Peptidase M10_N; 1.
DR PRINTS; PR00138; MATRXIN.
DR SMART; SM00120; HX; 4.

```

DR SMART: SM00235; ZnMc; 1.  
 DR PROSITE; PS00546; CYSTEINE SWITCH; 1.  
 DR PROSITE; PS00024; HEMOPEXIN; 1.  
 DR PROSITE; PS00142; ZINC\_PROTEASE; UNKNOWN\_1.  
 KW Metalloprotease; Protease.  
 SQ SEQUENCE 582 AA; 66068 MW; 390BCAA632D32DF8 CRC64;

Query Match 63.2%; Score 43; DB 2; Length 582;  
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 AAVALLPVLL 12

Db 539 AAVALLPVLL 548

RESULT 7

Q6AA98 PRELIMINARY; PRT; 671 AA.

ID O6AA98  
 AC O6AA98;  
 DT 25-OCT-2004 (TrEMBLrel. 28, Created)  
 DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)  
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)  
 DE Conserved protein containing thioredoxin domain.  
 GN OrderedLocusNames=PPA0563;  
 OS Propionibacterium acnes.  
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
 OC Propionibacteriaceae; Propionibacteriaceae; Propionibacterium.  
 OX NCBI\_TaxID=1747;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RC STRAIN=KPA171202 / DSM 16379;  
 RX PubMed=15286373; DOI=10.1126/science.1100330;  
 RA Bruggemann H., Henne A., Hofer P., Liesegang H., Wierer A.,  
 RA Strittmatter A., Hujer S., Duerre F., Gottschalk G.;  
 RT "The complete genome sequence of Propionibacterium acnes, a commensal  
 of human skin.";  
 RL Science 305:671-673(2004).  
 DR EMBL; AS017283; AAT82318.1; -.  
 DR InterPro; IPR004879; DUF255.  
 DR InterPro; IPR008928; Glyco trans\_6hp.  
 DR InterPro; IPR000834; Peptidase M14.  
 DR InterPro; IPR000169; Pept\_cys\_acsite.  
 DR Pfam; PF03190; DUF255; 1.  
 DR PROSITE; PS00133; CARBOXYPEPT\_ZN 2; UNKNOWN\_1.  
 DR PROSITE; PS00639; THIOL\_PROTEASE\_HIS; UNKNOWN\_1.  
 KW Complete proteome.  
 SQ SEQUENCE 671 AA; 74153 MW; 9354C3F7C2F9C544 CRC64;

Query Match 63.2%; Score 43; DB 2; Length 671;  
 Best Local Similarity 64.3%; Pred. No. 1.4e+02;  
 Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 KAAAVLLPVLLAAP 15

Db 555 KASVLLPVVSATP 568

RESULT 8

Q7W3J1 PRELIMINARY; PRT; 177 AA.

ID Q7W3J1  
 AC Q7W3J1;  
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)  
 DT 01-WAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Hypothetical protein.  
 GN OrderedLocusNames=BPP4042;  
 OS Bordetella parapertussis.  
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;  
 OC Alcaligenaceae; Bordetella.  
 OX NCBI\_TaxID=519;  
 RN [1]  
 RP SEQUENCE FROM N.A.

RC STRAIN=12822 / ATCC BAA-587;  
 RX MEDLINE=22827954; PubMed=12910271; DOI=10.1038/ng1227;  
 RA Parkhill J., Sebaihia M., Preston A., Murphy L.D., Thomson N.R.,  
 RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,  
 RA Cerdano-Tarraga A.-M., Temple L., James K.D., Harris B., Quail M.A.,  
 RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,  
 RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,  
 RA Feltwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,  
 RA Leather S., Moule S., Norberczak H., O'Neill S., Ormond D., Price C.,  
 RA Rabinowitz E., Rutter S., Sanders M., Saunders D., Seeger K.,  
 RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,  
 RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;  
 RT "Comparative analysis of the genome sequences of Bordetella pertussis,  
 Bordetella parapertussis and Bordetella bronchiseptica.";  
 RL Nat. Genet. 35:32-40(2003).  
 DR EMBL; BX640435; CAE39325.1; -.  
 DR InterPro; IPR003509; UPP0102.  
 DR Pfam; PF02021; UPP0102; 1.  
 DR TIGRFAMs; TIGR00252; UPP0102; 1.  
 KW Complete proteome.  
 SQ SEQUENCE 177 AA; 19268 MW; 2A1BDFD4964FFBBB CRC64;

Query Match 61.8%; Score 42; DB 2; Length 177;  
 Best Local Similarity 81.8%; Pred. No. 61;  
 Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 KAAAVLLPVLL 12

Db 115 RAALLLPVLL 125

RESULT 9

Q7WEW6 PRELIMINARY; PRT; 177 AA.

ID Q7WEW6  
 AC Q7WEW6;  
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Hypothetical protein.  
 GN OrderedLocusNames=BB4515;  
 OS Bordetella bronchiseptica (Alcaligenes bronchisepticus).  
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;  
 OC Alcaligenaceae; Bordetella.  
 OX NCBI\_TaxID=518;  
 RN [1]  
 RP SEQUENCE FROM N.A.

RC STRAIN=RB50 / ATCC BAA-588;  
 RX MEDLINE=22827954; PubMed=12910271; DOI=10.1038/ng1227;  
 RA Parkhill J., Sebaihia M., Preston A., Murphy L.D., Thomson N.R.,  
 RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,  
 RA Cerdano-Tarraga A.-M., Temple L., James K.D., Harris B., Quail M.A.,  
 RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,  
 RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,  
 RA Feltwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,  
 RA Leather S., Moule S., Norberczak H., O'Neill S., Ormond D., Price C.,  
 RA Rabinowitz E., Rutter S., Sanders M., Saunders D., Seeger K.,  
 RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,  
 RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;  
 RT "Comparative analysis of the genome sequences of Bordetella pertussis,  
 Bordetella parapertussis and Bordetella bronchiseptica.";  
 RL Nat. Genet. 35:32-40(2003).  
 DR EMBL; BX640450; CAE34878.1; -.  
 DR InterPro; IPR003509; UPP0102.  
 DR Pfam; PF02021; UPP0102; 1.  
 DR TIGRFAMs; TIGR00252; UPP0102; 1.  
 KW Complete proteome.  
 SQ SEQUENCE 177 AA; 19418 MW; BF6CD86DE3838146 CRC64;

Query Match 61.8%; Score 42; DB 2; Length 177;  
 Best Local Similarity 81.8%; Pred. No. 61;  
 Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 KAAAVLLPVLL 12



```

RA Nelson W.C., Ayodeji B., Kraul M., Shetty J., Malek J.A.,
RA Van Aken S.E., Riedmuller S., Tettelin H., Gill S.R., White O.,
RA Salzberg S.B., Hoover D.B., Lindler L.E., Halling S.M., Boyle S.M.,
RA Fraser C.M.;
RT "The Brucella suis genome reveals fundamental similarities between
RT animal and plant pathogens and symbionts.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:13148-13153 (2002).
DR EMBL; AF014478; AN30768.1; -.
DR TIGR; BR1873; -.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR000620; DUF6.
DR Pfam; PF00892; DUF6; 1.
KW Complete proteome.
SQ SEQUENCE 293 AA; 30733 MW; 4DFD4347112050AE CRC64;

Query Match 61.8%; Score 42; DB 2; Length 293;
Best Local Similarity 81.8%; Pred. No. 97;
Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 AAALLPVLAA 13
Db 156 AAALLPVLAA 166

RESULT 14
Q8YJ99 PRELIMINARY; PRT; 297 AA.
AC Q8YJ99
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE TRANSPORTER, DME FAMILY.
GN OrderedLocusNames=BMEI0187;
OS Brucella melitensis.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Brucellaceae; Brucella.
OX NCBI_TaxID=29459;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=16M / ATCC 23456 / Biotype 1;
RX MEDLINE=20020109; PubMed=11756688; DOI=10.1073/pnas.221575398;
RA DelVecchio V.G., Kapral V., Redkar R.J., Patra G., Mujar C., Los T.,
RA Ivanova N., Anderson I., Bhattacharya A., Lykidis A., Reznik G.,
RA Jablonski L., Larsen N., D'Souza M., Bernal A., Mazur M., Goltzman E.,
RA Selkov E., Elzer F.H., Hagius S., O'Callaghan D., Letesson J.-J.,
RA Haselkorn R., Kyrpides N.C., Overbeek R.;
RT "The genome sequence of the facultative intracellular pathogen
RT Brucella melitensis.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:443-448 (2002).
DR EMBL; AF009460; AAL51369.1; -.
DR FIR; AF3275; AF3275.
DR GO; GO:0016020; C:membrane; IEA.
DR Pfam; PF00892; DUF6; 1.
KW Complete proteome.
SQ SEQUENCE 297 AA; 31306 MW; E8CA2D3CEAB94B85 CRC64;

Query Match 61.8%; Score 42; DB 2; Length 297;
Best Local Similarity 81.8%; Pred. No. 98;
Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 AAALLPVLAA 13
Db 160 AAALLPVLAA 170

RESULT 15
Q7WNE5 PRELIMINARY; PRT; 323 AA.
AC Q7WNE5
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Putative exported protein.

```

```

GN OrderedLocusNames=BBI095;
OS Bordetella bronchiseptica (Alcaligenes bronchisepticus).
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
OX NCBI_TaxID=518;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RB50 / ATCC BAA-588;
RX MEDLINE=22827954; PubMed=12910271; DOI=10.1038/ngi227;
RA Parkhill J., Sebaihia M., Preston A., Murphy L.D., Thomson N.R.,
RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA Cerdano-Tarraga A.-M., Temple L., James K.D., Harris B., Quail M.A.,
RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA Felwell T., Goble A., Hamlin N., Hauser H., Holtroyd S., Jagsis K.,
RA Leather S., Moule S., Norberczak H., O'Neill S., Ormond D., Price C.,
RA Rabinowitsch E., Rutter S., Sanders M., Saunders D., Seeger K.,
RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
RA Unwin L., Whitehead S., Barrell B.G., Maskell D.J.;
RT "Comparative analysis of the genome sequences of Bordetella pertussis,
RT Bordetella parapertussis and Bordetella bronchiseptica.";
RL Nat. Genet. 35:32-40 (2003).
DR EMBL; BX640440; CAE31593.1; -.
DR GO; GO:0030288; C:periplasmic space (sensu Gram-negative Bact. . .; IEA.
DR InterPro; IPR005064; UPF0065.
DR Pfam; PF03401; Bug; 1.
KW Complete proteome.
SQ SEQUENCE 323 AA; 34109 MW; 9AD672EE4BF646E4 CRC64;

Query Match 61.8%; Score 42; DB 2; Length 323;
Best Local Similarity 60.0%; Pred. No. 1.1e+02;
Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 KKAARVLPVLLAAP 15
Db 2 KRYAGILLGMLLAAP 16

RESULT 16
Q89HQ1 PRELIMINARY; PRT; 422 AA.
AC Q89HQ1
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Blr5939 protein.
GN OrderedLocusNames=blr5939;
OS Bradyrhizobium japonicum.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Bradyrhizobium.
OX NCBI_TaxID=375;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=USDA110;
RX MEDLINE=22484998; PubMed=12597275;
RA Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiumi T.,
RA Sasamoto S., Watanabe A., Idesawa K., Iriguchi M., Kawashima K.,
RA Kohara M., Matsumoto M., Shimpo S., Tsuruoka H., Wada T., Yamada M.,
RA Tabata S.;
RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium
RT Bradyrhizobium japonicum USDA110.";
RL DNA Res. 9:189-197 (2002).
DR EMBL; AP005956; BAC51204.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005215; P:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR010645; DUF1228.
DR InterPro; IPR007114; MFS.
DR Pfam; PF06779; DUF1228; 1.
DR PROSITE; PS50850; MFS; 1.
KW Complete proteome.
SQ SEQUENCE 422 AA; 43914 MW; 2374A14B68E31761 CRC64;

```



```

Query Match      61.8%; Score 42; DB 2; Length 422;
Best Local Similarity 64.3%; Pred. No. 1.4e+02;
Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 KAAAVLLPVLLAAP 15
   :|||||:
Db 311 EAAVLLPTVSTAP 324

RESULT 17
O49467 PRELIMINARY; PRT; 939 AA.
AC O49467;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein F24J7.50 (Hypothetical protein AT4g19490).
GN Name=F24J7.50; Synonyms=AT4g19490;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsia.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Bevan M., Vitale D., Liguori R., Argirou A., De Simone V.,
RA Bancroft I., Mewes H.W., Mayer K.F.X., Lemcke K., Schueller C.;
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Vitale D., Liguori R., Argirou A., De Simone V., Mewes H.W.,
RA Lemcke K., Mayer K.F.X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL021768; CAA16926.1; -
DR EMBL; AL161551; CAB78951.1; -
DR PIR; T05209; T05209.
KW Hypothetical protein.
SQ SEQUENCE 939 AA; 102496 MW; 9DAEBA3F6105CEEC CRC64;

Query Match      61.8%; Score 42; DB 2; Length 939;
Best Local Similarity 69.2%; Pred. No. 2.8e+02;
Matches 9; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 KKAARVLLPVLLA 13
   |||||:|:|:|:|
Db 424 KKAADLLPILVA 436

RESULT 18
CU65_LOCMI STANDARD; PRT; 145 AA.
AC P45585;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Cuticle protein 65 (LM-65) (LM-ACP 65).
OS Locusta migratoria (Migratory locust).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Orthopteroidea; Orthoptera; Caelifera; Acridomorpha;
OC Acridoidea; Acrididae; Oedipodinae; Locusta.
OX NCBI_TaxID=7004;
RN [1]
RP SEQUENCE, AND MASS SPECTROMETRY.
RC TISSUE=Wing;
RX MEDLINE=95291280; PubMed=7773253; DOI=10.1016/0965-1748(94)00056-N;

```

```

RA Krogh T.N., Skou L., Roepstorff P., Andersen S.O., Højrup P.;
RT "Primary structure of proteins from the wing cuticle of the migratory
RT locust, Locusta migratoria."
RL Insect Biochem. Mol. Biol. 25:319-329(1995).
CC -!- FUNCTION: Component of the cuticle of migratory locust which
CC contains more than 100 different structural proteins.
CC -!- DOMAIN: The tetrapeptide (A-A-P-[AV]) repeats found throughout the
CC protein are also present in many proteins constituting the
CC protective envelope of other species.
CC -!- MASS SPECTROMETRY: MW=13319.9; METHOD=Electrospray; RANGE=1-145;
CC NOTE=Ref.1.
KW Cuticle; Direct protein sequencing; Repeat; Structural protein.
FT REPEAT 27 30 1.
FT REPEAT 33 37 2.
FT REPEAT 39 42 3.
FT REPEAT 86 89 4.
FT REPEAT 92 95 5.
FT REPEAT 98 101 6.
FT REPEAT 123 126 7.
SQ SEQUENCE 145 AA; 13318 MW; 6677E36A4D52148A CRC64;

Query Match      60.3%; Score 41; DB 1; Length 145;
Best Local Similarity 64.3%; Pred. No. 75;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 KAAAVLLPVLLAAP 15
   |||||:|:|:|
Db 81 KVAAPVAVAP 94

RESULT 19
Q7PS92 PRELIMINARY; PRT; 195 AA.
AC Q7PS92;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE ENSANGP0000018919.
GN Name=ENSANGP0000018919.
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Anopheles.
OX NCBI_TaxID=180454;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PEST;
RA Anopheles Genome Sequencing Consortium;
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAB0100844; EAA05987.2; -.
DR InterPro; IPR001969; Pept_Asp_AS.
DR PROSITE; PS00141; ASP_PROTEASE; UNKNOWN 1.
SQ SEQUENCE 195 AA; 20951 MW; 3DF52FD0418AF6A7 CRC64;

Query Match      60.3%; Score 41; DB 2; Length 195;
Best Local Similarity 76.9%; Pred. No. 98;
Matches 10; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 AAVALLPVLLAAP 15
   :|||||:
Db 13 SAAVLLQPLLAAP 25

RESULT 20
Q9BIH2 PRELIMINARY; PRT; 195 AA.
AC Q9BIH2;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein GSG10 precursor.

```

RL	Submitted (JUL-2004) to the EMBL/GenBank/DBDJ databases.
DR	EMBL; CR382125; CAG99157.1; -.
DR	InterPro; IPR000504; RNA_rec_mot.
DR	DR PFam; PFO0076; RRM 1; 1.
DR	SMART; SM00360; RRM; 1.
DR	PROSITE; PS0102; RRM; 1.
SQ	SEQUENCE 222 AA; 24084 MW; 1CAC420F75DA49EC CRC64;
Query Match	60.3%; Score 41; DB 2; Length 222;
Best Local Similarity	69.2%; Pred. No. 1.1e+02;
Matches 11; Conservative	1; Mismatches 3; Indels 2;
QY	1 KKA--AAVLLPVLLAAP 15          :
Dd	49 KKAASAAALVPALLRAP 65            :
RESULT 22	
ID Q7U481	PRELIMINARY; PRT; 231 AA.
AC Q7U481;	
DT 01-OCT-2003 (TrEMBLrel. 25, Created)	
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)	
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)	
DE Hypothetical protein.	
GN OrderedLocusNames=SYNMW2190;	
OS Synchococcus sp. (strain WH8102).	
OC Bacteria; Cyanobacteria; Chroococcales; Synechococcus.	
OX NCBI_TaxID=84588;	
RN [1]	
RX SEQUENCE FROM N.A.	
RY MEDLINE=22825697; PubMed=12917641; DOI=10.1038/nature01943;	
RA Palenik B., Brahmasha B., Laximer F.W., Land M.L., Hauser L.,	
RA Chain P., Lamerdin J.E., Regala W., Allen E.E., McCarren J.,	
RA Paulsen I.T., Dufresne A., Bartensky F., Webb E.A., Waterbury	
RT "The genome of a motile marine Synechococcus";	
RT Nature 424:1037-1042(2003).	
DR EMBL; BX569694; CAB08705.1; -.	
KW Complete proteome; Hypothetical protein.	
SQ SEQUENCE 231 AA; 24714 MW; 63025D8FA9A7FC2 CRC64;	
Query Match	60.3%; Score 41; DB 2; Length 231;
Best Local Similarity	69.2%; Pred. No. 1.1e+02;
Matches 9; Conservative	1; Mismatches 3; Indels 0;
QY	3 AAAVLLPVLLAAP 15          :
Dd	24 AAHVLPVFQQAP 36            :
RESULT 23	
Q82NQ4	PRELIMINARY; PRT; 659 AA.
ID Q82NQ4	
AC Q82NQ4;	
DT 01-JUN-2003 (TrEMBLrel. 24, Created)	
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)	
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)	
DE Putative secreted amidase.	
GN OrderedLocusNames=SAV1238;	
OS Streptomyces avermitilis.	
OC Bacteria; Actinobacteria; Actinomycetales; Actinomycetaceae; Streptomycinae; Streptomyces.	
OX NCBI_TaxID=33903;	
RN [1]	
RP SEQUENCE FROM N.A.	
RC STRAIN=MA-4680;	
RX MEDLINE=21477403; PubMed=11572948; DOI=10.1073/pnas.211433198;	
RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,	
RA Shinoue M., Takahaehi Y., Horikawa H., Nakazawa H., Osone T.,	
RA Kikuchi H., Shiba T., Sakai Y., Hattori M.	
RT "Genome sequence of an industrial microorganism Streptomyces	
RT avermitilis: deducing the ability of producing secondary	
RT metabolites.";	

```

RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
RN [2]
RN SEQUENCE FROM N.A.
RC STRAIN=MA-4680;
RX MEDLINE=22608306; PubMed=12692562;
RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
RA Sakaki Y., Hattori M., Omura S.;
RT "Complete genome sequence and comparative analysis of the industrial
RT microorganism Streptomyces avermitilis.";
RL Nat. Biotechnol. 21:526-531(2003).
DR EMBL; AP050526; BAC68948.1; -.
DR GO; GO:0008745; F:N-acetylmuramoyl-L-alanine amidase activity; IEA.
DR GO; GO:0009253; P:peptidoglycan catabolism; IEA.
DR InterPro; IPR002502; Amidase 2.
DR Pfam; PF01510; Amidase 2; 1.
DR SMART; SM00644; Ami_2; 1.
KW Complete proteome.
SQ SEQUENCE 659 AA; 70218 MW; BA3905045D2792FF CRC64;

Query Match 60.3%; Score 41; DB 2; Length 659;
Best Local Similarity 69.2%; Pred. No. 3e+02;
Matches 9; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

Qy 3 AAALLPVLAAAP 15
Db 26 SAALLPLLGAP 38

RESULT 24
QY98Y4
ID QY98Y4 PRELIMINARY; PRT; 1172 AA.
AC QY98Y4;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE 1172aa long hypothetical antibiotic transport-associated protein.
GN OrderedLocustNames=ABE2500;
OS Aeropyrum pernix.
OC Archaea; Crenarchaeota; Thermoprotei; Desulfurococcaceae;
OC Desulfurococcaceae; Aeropyrum.
OX NCBI_TaxID=56636;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K1;
RX MEDLINE=93310339; PubMed=10382966;
RA Kwarabayasi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y.,
RA Jin-no K., Takahashi M., Sekine M., Baba S.-I., Ankai A., Kosugi H.,
RA Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
RA Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,
RA Yamazaki J., Kusbida N., Oguchi A., Aoki K.-I., Kubota K.,
RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.;
RT "Complete genome sequence of an aerobic hyper-thermophilic
RT crenarchaeon, Aeropyrum pernix K1.";
RL DNA Res. 6:83-101(1999).
DR EMBL; AP000064; BAA81516.1; -.
DR PIR; D72482; D72482
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 1172 AA; 124325 MW; 974421D7A5470D5 CRC64;

Query Match 60.3%; Score 41; DB 2; Length 1172;
Best Local Similarity 69.2%; Pred. No. 5e+02;
Matches 9; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 2 KAAAVLLPVLAA 14
Db 848 KAAPVLLPLVVA 860

RESULT 25
PSBY SYNEL
ID PSBY SYNEL STANDARD; PRT; 43 AA.
AC Q8DKM3;
DT 10-OCT-2003 (Rel. 42, Created)

```

```

DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Photosystem II protein Y.
GN Name=psby; OrderedLocustNames=ts10836;
OS Synechococcus elongatus (Thermosynechococcus elongatus).
OC Bacteria; Cyanobacteria; Chroococcales; Synechococcus.
OX NCBI_TaxID=32046;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=BP-1;
RX MEDLINE=22225144; PubMed=12240834;
RA Nakamura Y., Kaneo T., Sato S., Ikeuchi M., Katoh H., Sasamoto S.,
RA Watanabe A., Iriguchi M., Kawashima K., Kimura T., Kishida Y.,
RA Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., Nakazaki N.,
RA Shimo S., Sugimoto M., Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the thermophilic cyanobacterium
RT Thermosynechococcus elongatus BP-1.";
RL DNA Res. 9:123-130(2002).
CC -!- FUNCTION: Manganese-binding polypeptide with L-arginine
CC metabolizing enzyme activity. Component of the core of photosystem
CC II (By similarity).
CC -!- SUBCELLULAR LOCATION: Cellular thylakoid membrane.
CC -!- SIMILARITY: Belongs to the psby family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AP005371; BAC08388.1; -.
CC HANAP; MF_00717; -.
CC InterPro; IPR009388; Psby.
CC Pfam; PF06298; Psby; 1.
CC Complete proteome; Photosynthesis; Photosystem II; Thylakoid;
CC Transmembrane.
KW TRANSMEM 7 25 Potential.
SQ SEQUENCE 43 AA; 4773 MW; B12F8F16F2EDF86 CRC64;

Query Match 58.8%; Score 40; DB 1; Length 43;
Best Local Similarity 69.2%; Pred. No. 36;
Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 2 KAAAVLLPVLAA 14
Db 6 RVLVLLPVLAA 18

RESULT 26
Q6GX97
ID Q6GX97 PRELIMINARY; PRT; 67 AA.
AC Q6GX97;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Matrix metalloproteinase 14 (Fragment).
GN Name=Mmp14;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=BALB/c; TISSUE=Uterus;
RA Ulrich R., Gerhauser I., Seeliger F., Baumgaertner W., Alldinger S.;
RT "Expression profile of matrix-metalloproteinases and their inhibitors
RT in acute and chronic demyelinating theiler's murine
RT encephalomyelitis.";
RL Acta Neuropathol. 108:366-366(2004).
RN [2]
RN SEQUENCE FROM N.A.

```

RC STRAIN=BALB/c; TISSUE=Uterus;  
 RA Ulrich R.G., Seeliger F., Alldinger S., Baumgaertner W.;  
 RL Submitted (MAY-2004) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AY622974; AAT46405.1; -.  
 FT NON\_TER 1  
 FT NON\_TER 67  
 SQ SEQUENCE 67 AA; 7104 MW; 52CDDDF3957E3FBA CRC64;

Query Match 58.8%; Score 40; DB 2; Length 67;  
 Best Local Similarity 90.0%; Pred. No. 54;  
 Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 AAAYLPLVLL 12  
 |||||  
 Db 53 AAAYLPLVLL 62

RESULT 27  
 ID NUTP SPIOL STANDARD; PRT; 117 AA.  
 AC P10976;  
 DT 01-JUL-1989 (Rel. 11, Created)  
 DT 01-AUG-1992 (Rel. 23, Last sequence update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE Nonspecific lipid-transfer protein precursor (LTP) (Phospholipid transfer protein) (PPTP).  
 OS Spinacia oleracea (Spinach).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Caryophyllales; Amaranthaceae; Spinacia.  
 OX NCBI\_TaxID=3562;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Bernhard W.R., Thoma S., Botella J., Somerville C.R.;  
 RT "Isolation of a cDNA clone for spinach lipid transfer protein and evidence that the protein is synthesized by the secretory pathway.";  
 RL Plant Physiol. 95:164-170(1991).  
 RN [2]  
 RP SEQUENCE OF 27-117.

RC TISSUE=Leaf;  
 RX MEDLINE=87275922; PubMed=3609015;  
 RA Bouillon P., Drischel C., Vergnolle C., Duranton H., Kader J.-C.;  
 RT "The primary structure of spinach-leaf phospholipid-transfer protein.";  
 RL Eur. J. Biochem. 166:387-391(1987).  
 CC -1- FUNCTION: Plant nonspecific lipid-transfer proteins transfer phospholipids as well as galactolipids across membranes. May play a role in wax or cutin deposition in the cell walls of expanding epidermal cells and certain secretory tissues.  
 CC -1- SIMILARITY: Belongs to the plant LTP family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; M58635; AAA34032.1; -.  
 DR FIC; T09155; T09155.  
 DR HSSP; P19656; 1FK5.  
 DR InterPro; IPR003612; AAI.  
 DR InterPro; IPR000528; Plant\_LTP.  
 DR Pfam; PF00234; Tryp\_alpha\_aml1; 1.  
 DR PRINTS; PR00382; LIPIDTRANSF.  
 DR SMART; SM00499; AAI; 1.  
 DR PROSITE; PS00597; PLANT\_LTP; 1.  
 KW Direct protein sequencing; Lipid-binding; Signal; Transport.  
 FT SIGNAL 1 26  
 FT CHAIN 27 117 Nonspecific lipid-transfer protein.  
 FT DISULFID 30 76 By similarity.  
 FT DISULFID 40 53 By similarity.

FT DISULFID 54 99 By similarity.  
 FT DISULFID 74 113 By similarity.  
 FT CONFLICT 53 54 CC -> SS (in Ref. 2).  
 SQ SEQUENCE 117 AA; 11425 MW; 19F969F94D58FA34 CRC64;

Query Match 58.8%; Score 40; DB 1; Length 117;  
 Best Local Similarity 60.0%; Pred. No. 91;  
 Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 KKAAYLPLVLLAAP 15  
 |||||  
 Db 8 KLAAYLPLVLLAAP 22

RESULT 28  
 Q854U6 PRELIMINARY; PRT; 138 AA.  
 ID Q854U6;  
 AC Q854U6;  
 DT 01-JUN-2003 (TREMBLrel. 24, Created)  
 DT 01-JUN-2003 (TREMBLrel. 24, Last sequence update)  
 DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)  
 DE Gp54.  
 OS Mycobacteriophage Che9c.  
 OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae.  
 OX NCBI\_TaxID=205872;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=2592660; PubMed=12705866; DOI=10.1016/S0092-8674(03)00233-2;  
 RA Pedulla M.L., Ford M.E., Houtz J.M., Karthikeyan T., Wadsworth C., Lewis J.A., Jacobs-Sera J., Falbo J., Gross J., Panunzio N.R., Brucker W., Kumar V., Kandasamy J., Keenan L., Bardarov S., Kriakov J., Lawrence J.G., Jacobs W.R. Jr., Hendrix R.W., Hatfull G.F.;  
 RA "Origins of highly mosaic mycobacteriophage genomes.";  
 RT Cell 113:171-182(2003).  
 RL EMBL; AY129333; AAN12612.1; -.  
 SQ SEQUENCE 138 AA; 14405 MW; 120D21C12BBADAF1 CRC64;

Query Match 58.8%; Score 40; DB 2; Length 138;  
 Best Local Similarity 69.2%; Pred. No. 11e+02;  
 Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 3 AAAYLPLVLLAAP 15  
 |||||  
 Db 45 AAAYLPLVLLAAP 57

RESULT 29  
 Q6W1Z2 PRELIMINARY; PRT; 138 AA.  
 ID Q6W1Z2;  
 AC Q6W1Z2;  
 DT 05-JUL-2004 (TREMBLrel. 27, Created)  
 DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)  
 DE Hypothetical protein.  
 GN ORFNames=NGR00204;  
 OS Rhizobium sp. (strain NGR234).  
 OG Plasmid megaplasmid 2.  
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
 OC Rhizobiaceae; Rhizobium/Agrobacterium group; Rhizobium.  
 OX NCBI\_TaxID=394;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=NGR234;  
 RX PubMed=14702322; DOI=10.1128/JB.186.2.535-542.2004;  
 RA Streit W.R., Schmitz R.A., Perret X., Staehelin C., Deakin W.J., Raasch C., Liesegang H., Broughton W.J.;  
 RT "An evolutionary hot spot: the pNGR234b replicon of Rhizobium sp. strain NGR234.";  
 RL J. Bacteriol. 186:535-542(2004).  
 DR EMBL; AY316746; AAQ87226.1; -.  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR InterPro; IPR011592; Surf4\_rel.

```

DR ProDom; PD010195; Surf4_rel; 1.
KW Hypothetical protein; Plasmid.
SQ SEQUENCE 138 AA; 14245 MW; C87156ED1EB080F7 CRC64;

Query Match 58.8%; Score 40; DB 2; Length 138;
Best Local Similarity 75.0%; Pred. No. 1.1e+02;
Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 3 AAVALLPVLLAA 14
Db 75 AALVLPVLLGA 86

RESULT 30
CU64 LOCMI STANDARD; PRT; 152 AA.
ID P11739
AC P11739
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Cuticle protein 64 (LM-64) (LM-ACP 64).
OS Locusta migratoria (Migratory locust).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Orthopteroidea; Orthoptera; Caelifera; Acridomorpha;
OC Acridoidea; Acrididae; Oedipodinae; Locusta.
OX NCBI_TaxID=7004;
RN [1]
RP SEQUENCE, AND MASS SPECTROMETRY.
RC TISSUE=wing;
RX MEDLINE=95291280; PubMed=7773253; DOI=10.1016/0965-1748(94)00056-N;
RA Krogh T.N., Skou L., Roepstorff P., Andersen S.O., Hoegstrup P.;
RT "Primary structure of proteins from the wing cuticle of the migratory
RT locust, Locusta migratoria.";
RL Insect Biochem. Mol. Biol. 25:319-329 (1995).
RN [2]
RP SEQUENCE OF 1-59.
RX MEDLINE=86108304; PubMed=3943519;
RA Hoegstrup P., Andersen S.O., Roepstorff P.;
RT "Isolation, characterization, and N-terminal sequence studies of
RT cuticular proteins from the migratory locust, Locusta migratoria.";
RL Eur. J. Biochem. 154:153-159 (1986).
CC -!- FUNCTION: Component of the cuticle of migratory locust which
CC contains more than 100 different structural proteins.
CC -!- DOMAIN: The tetrapeptide (A-A-P-[AV]) repeats found throughout the
CC protein are also present in many proteins constituting the
CC protective envelope of other species.
CC -!- MASS SPECTROMETRY: MW=14050.8; METHOD=Electrospray; RANGE=1-152;
CC NOTE=Ref.1.
KW Cuticle; Direct protein sequencing; Repeat; Structural protein.
FT REPEAT 27 30 1.
FT REPEAT 33 37 2.
FT REPEAT 39 42 3.
FT REPEAT 86 89 4.
FT REPEAT 92 95 5.
FT REPEAT 98 101 6.
FT REPEAT 127 130 7.
FT CONFLICT 11 11 HG -> S (in Ref. 2).
FT CONFLICT 21 22 HG -> PH (in Ref. 2).
FT CONFLICT 27 27 A -> Y (in Ref. 2).
FT CONFLICT 30 32 AAI -> LAL (in Ref. 2).
FT CONFLICT 34 34 A -> Y (in Ref. 2).
FT CONFLICT 45 45 H -> Q (in Ref. 2).
SQ SEQUENCE 152 AA; 14050 MW; 0C127C22F28E2941 CRC64;

Query Match 58.8%; Score 40; DB 1; Length 152;
Best Local Similarity 57.1%; Pred. No. 1.2e+02;
Matches 8; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 2 KAAAVLLPVLLAAP 15
Db 81 KVAATAAPVAAP 94

```

```

RESULT 31
Q92KQ6 PRELIMINARY; PRT; 153 AA.
AC Q92KQ6
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical protein SMC02178.
GN ORFNames=SMC02178;
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1021;
RX MEDLINE=21396507; PubMed=11481430; DOI=10.1073/pnas.161294398;
RA Capela D., Barloy-Hubler F., Gouzy J., Bothe G., Ampe F., Batut J.,
RA Boistard P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,
RA Godrie T., Goffeau A., Kahn D., Kiss E., Lelaure V., Masuy D.,
RA Pohl T., Portetelle D., Puhler A., Purnelle B., Rampsperger U.,
RA Renard C., Thebault P., Vandenbol M., Weidner S., Galibert F.;
RT "Analysis of the chromosome sequence of the legume symbiont
RT Sinorhizobium meliloti strain 1021.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882 (2001).
DR EMBL; AL591783; CAC41906.1;
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 153 AA; 15998 MW; 604679A394F98672 CRC64;

Query Match 58.8%; Score 40; DB 2; Length 153;
Best Local Similarity 61.5%; Pred. No. 1.2e+02;
Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 3 AAVALLPVLLAAP 15
Db 14 AAVALPLVATAP 26

RESULT 32
Q8UDX0 PRELIMINARY; PRT; 158 AA.
AC Q8UDX0
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein Atul998.
GN OrderedLocustNames=Atul998;
OS Agrobacterium tumefaciens (strain CS8 / ATCC 33970).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
OX NCBI_TaxID=176299;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Dupont;
RX MEDLINE=21608550; PubMed=11743193; DOI=10.1126/science.1066804;
RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,
RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Bovee D. St.,
RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,
RA Kuyavlin T., Levy R., Li M.-J., McClelland E., Palmieri A.,
RA Raymond C., Rouse G., Saenphimmachak C., Wu Z., Romero P., Gordon D.,
RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
RA Chumley P., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
RA Nester E.W.;
RT "The genome of the natural genetic engineer Agrobacterium tumefaciens
RT CS8.";
RL Science 294:2317-2323 (2001).
DR EMBL; AE009151; AAL42991.1;
DR PIR; A97600; A97600.
DR PIR; A12821; A12821.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 158 AA; 17553 MW; 242AFA4D18DDF415 CRC64;

```

```
Query Match 58.8%; Score 40; DB 2; Length 158;
Best Local Similarity 69.2%; Pred. No. 1.2e+02;
Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 AAVALPVLAAAP 15
||| |||||
DB 11 AALALAPALLAAP 23

RESULT 33
Q7CY40 PRELIMINARY; PRT; 163 AA.
AC Q7CY40;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE AGR_C 3631p.
OS OrderedLocusNames=AGR_C 3631;
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
OX NCBI_TaxID=176299;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Cereon;
RX MEDLINE=21608551; PubMed=11743194; DOI=10.1126/science.1066803;
RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,
RA Quorillo B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,
RA Houmel K., Gordon J., Vaudin M., Iartchouk O., Epp A., Liu F.,
RA Wollam C., Allinger M., Doughty D., Scott C., Lappas C., Markelz B.,
RA Flanagan C., Crowell C., Gursen J., Lomo C., Sear C., Strub G.,
RA Cielo C., Slater S.;
RA "Genome sequence of the plant pathogen and biotechnology agent
RT Agrobacterium tumefaciens C58."
RL Science 294:2323-2328(2001).
DR EMBL; A5008117; AAK87754.1; -.
SQ SEQUENCE 163 AA; 18108 MW; 3A5AE6A29F2EE43B CRC64;

Query Match 58.8%; Score 40; DB 2; Length 163;
Best Local Similarity 69.2%; Pred. No. 1.2e+02;
Matches 9; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 3 AAVALPVLAAAP 15
||| |||||
DB 16 AALALAPALLAAP 28

RESULT 34
Q7S0K8 PRELIMINARY; PRT; 168 AA.
AC Q7S0K8;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Predicted protein.
GN Name=NCU09978.1;
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OK74A;
RA Galagan J.E., Calvo S.E., Borkovich K.A., Selker E.U., Read N.D.,
RA Jaffe D., Fitzhugh W., Ma L.-J., Smirnov S., Purcell S., Rehman B.,
RA Elkins T., Engels R., Wang S., Nielsen C.B., Butler J., Endrizzi M.,
RA Qui D., Ianakiev P., Pedersen D., Nelson M., Washburne M.,
RA Selitrenikoff C.P., Kinsey J.A., Braun E.L., Zelter A., Schulte U.,
RA Kothe G.O., Jedd G., Mewes W., Staben C., Marcotte E., Greenberg D.,
RA Roy A., Foley K., Naylor T., Thomann N., Barrett R., Gnerre S.,
RA Kamal M., Kamysysselis M., Mauceli E., Bielke C., Rudd S., Frishman D.,
RA Krystofova S., Rasmussen C., Metznerberg R.L., Perkins D.D., Kroken S.,
```

---

```
RA Cogoni C., Macino G., Catcheside D., Li W., Pratt R.J., Osmari S.A.,
RA DeSouza C.C., Glass L., Orbach M.J., Berglund J., Voelker R.,
RA Yarden O., Pismam M., Seiler S., Dunlap J., Radford A., Aramayo R.,
RA Natvig D.O., Alex L.A., Mannhaupt G., Ebbole D.J., Freitag M.,
RA Paulsen I., Sachs M.S., Lander E.S., Nussbaum C., Birren B.;
RT "The Genome Sequence of the Filamentous Fungus Neurospora crassa.";
RL Nature 0:0-0(2003).
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AABX01000491; EAA28860.1; -.
SQ SEQUENCE 168 AA; 17345 MW; 9345611183ECA1D8 CRC64;

Query Match 58.8%; Score 40; DB 2; Length 168;
Best Local Similarity 53.3%; Pred. No. 1.3e+02;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 KKAAVLLPVLAAAP 15
||| |||||
DB 43 QKAGIALPPLLLSP 57

RESULT 35
Q9HIX9 PRELIMINARY; PRT; 252 AA.
AC Q9HIX9;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein Tal197.
GN OrderedLocusNames=Tal197;
OS Thermoplasma acidophilum.
OC Archaea; Euryarchaeota; Thermoplasmatia; Thermoplasmatales;
OC Thermoplasmataceae; Thermoplasma.
OX NCBI_TaxID=2303;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DSM 1728;
RX MEDLINE=20479972; PubMed=11029001; DOI=10.1038/35035069;
RA Ruepp A., Graml W., Santos-Martinez M.-L., Koretke K.K., Volker C.,
RA Mewes H.-W., Frishman D., Stocker S., Lupas A.N., Baumeister W.;
RT "The genome sequence of the thermoacidophilic scavenger Thermoplasma
RT acidophilum."
RL Nature 407:508-513(2000).
DR EMBL; AL445066; CAC12322.1; -.
DR InterPro; IPR009058; Wing_hlx_DNA_bnd.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 252 AA; 27640 MW; BD2916EC4C4B9F06 CRC64;

Query Match 58.8%; Score 40; DB 2; Length 252;
Best Local Similarity 71.4%; Pred. No. 1.8e+02;
Matches 10; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 KKAAVLLPVLAAAP 14
||| |||||
DB 169 KKAADLSPALFAA 182

RESULT 36
Q89D56 PRELIMINARY; PRT; 279 AA.
AC Q89D56;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE B1r7587 protein.
GN OrderedLocusNames=b1r7587;
OS Bradyrhizobium japonicum.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Bradyrhizobium.
OX NCBI_TaxID=375;
RN [1]
RP SEQUENCE FROM N.A.
```

```

RC STRAIN-USDA110;
RX MEDLINE=22484998; PubMed=12597275;
RA Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiumi T.,
RA Sasamoto S., Watanabe A., Ideesawa K., Iriguchi M., Kawashima K.,
RA Kohara M., Matsumoto M., Shimpo S., Tsuruoka H., Wada T., Yamada M.,
RA Tabata S.;
RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium
RT Bradyrhizobium japonicum USDA110.";
RL DNA Res. 9:189-197(2002).
DR EMBL; AP005962; BAC52852.1; -.
KW Complete proteome.
SQ SEQUENCE 279 AA; 29666 MW; 27651945C0E13C39 CRC64;

Query Match 58.8%; Score 40; DB 2; Length 279;
Best Local Similarity 58.3%; Pred. No. 2e+02;
Matches 7; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 4 KAAVLLPVLLAAP 15
|||:|:|:|
Db 115 KAAVLLPVIVGTP 126

RESULT 37
Q7QB91 PRELIMINARY; PRT; 297 AA.
AC Q7QB91;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE AGCP2447.
GN Name=agCG45893; ORFNames=ENSGG000000017897;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Anopheles.
OX NCBI_TaxID=180454;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PEST;
RA Anopheles Genome Sequencing Consortium;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAB01008890; EAA08668.1; -.
SQ SEQUENCE 297 AA; 30551 MW; 7D60098FFB259BB5 CRC64;

Query Match 58.8%; Score 40; DB 2; Length 297;
Best Local Similarity 69.2%; Pred. No. 2.1e+02;
Matches 9; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KKAARAVLLPVLLA 13
||||:|:|:|
Db 174 KKAARKILGPLLLA 186

RESULT 38
Q83R77 PRELIMINARY; PRT; 298 AA.
ID Q83R77
AC Q83R77;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Porin OmpN.
GN Name=ompN; OrderedLocusNames=SF1823;
OS Shigella flexneri.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Shigella.
OX NCBI_TaxID=623;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=301 / Serotype 2a;
RX MEDLINE=22272406; PubMed=12384590; DOI=10.1093/nar/gkf566;
RA Jin Q., Yuan Z., Xu J., Wang Y., Lu W., Wang J., Liu H.,

```

```

RA Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
RA Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Qian B., Ding K., Chen S.,
RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
RA Yu J.;
RT "Genome sequence of Shigella flexneri 2a: insights into pathogenicity
RT through comparison with genomes of Escherichia coli K12 and O157.";
RL Nucleic Acids Res. 30:4432-4441(2002).
DR EMBL; AE015202; AAN43388.1; -.
DR HSSP; P02931; 1GFN.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019867; C:outer membrane; IEA.
DR GO; GO:0015288; F:porin activity; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR001897; Porin_bac.
DR InterPro; IPR001702; Porin_Gram-ve.
DR InterPro; IPR000408; Reg_chrom_condens.
DR Pfam; PF00267; Porin_1; 1.
DR PRINTS; PR00183; ECOLIPORIN.
DR PRINTS; PR00182; ECOLNEIPORIN.
DR PROSITE; PS00626; RCC1_2; UNKNOWN_1.
KW Complete proteome; Porin.
SQ SEQUENCE 298 AA; 32637 MW; EE86B49FBD32683E CRC64;

Query Match 58.8%; Score 40; DB 2; Length 298;
Best Local Similarity 57.1%; Pred. No. 2.1e+02;
Matches 8; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 KKAARAVLLPVLLA 14
:|:|:|:|
Db 3 RKVLALLIPALLAA 16

RESULT 39
Q9YDB2 PRELIMINARY; PRT; 390 AA.
ID Q9YDB2
AC Q9YDB2;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein APE1001.
GN OrderedLocusNames=APE1001;
OS Aeropyrum pernix.
OC Archaea; Crenarchaeota; Thermoprotei; Desulfurococcales;
OC Desulfurococaceae; Aeropyrum.
OX NCBI_TaxID=56636;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K1;
RX MEDLINE=99310139; PubMed=10382966;
RA Kavarabayasi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y.,
RA Jin-no K., Takahashi M., Sekine M., Baba S.-I., Ankai A., Koeugi H.,
RA Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
RA Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,
RA Yamazaki J., Kishida N., Oguchi A., Aoki K.-I., Kubota K.,
RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.;
RT "Complete genome sequence of an aerobic hyper-thermophilic
RT Crenarchaeon, Aeropyrum pernix K1.";
RL DNA Res. 6:83-101(1999).
DR EMBL; AP000060; BAA79985.1; -.
DR PIR; A72698; A72698.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 390 AA; 41090 MW; 9E98D66EAEFD207 CRC64;

Query Match 58.8%; Score 40; DB 2; Length 390;
Best Local Similarity 75.0%; Pred. No. 2.7e+02;
Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 4 KAAVLLPVLLAAP 15
||||:|:|:|
Db 180 KAAVLLPVLLAAP 191

```

```

RESULT 40
Q8U119 PRELIMINARY; PRT; 431 AA.
AC Q8U119;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Exopolysaccharide production protein.
GN Name=exQ; OrderedLocusNames=Atu0481;
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
OX NCBI_TaxID=176299;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Dupont;
RX MEDLINE=21608550; PubMed=11743193; DOI=10.1126/science.1066804;
RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,
RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Bovee D. Sr.,
RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,
RA Kutyavin T., Levy R., Li M.-J., McClelland E., Palmieri A., Gordon D.,
RA Raymond C., Rouse G., Saenphimmachak C., Wu Z., Romero P., Gordon D.,
RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
RA Chumley P., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
RA Nester E.W.;
RT "The genome of the natural genetic engineer Agrobacterium tumefaciens
RT C58."
RL Science 294:2317-2323(2001).
DR EMBL; AE009018; AAL41500.1; -.
DR PIR; AF2635; AF2635.
DR PIR; E97417; E97417.
DR Pfam; PF04932; Wzy_C; 1.
DR PROSITE; PS00213; LIPOCALIN; UNKNOWN_1.
SQ SEQUENCE 431 AA; 47698 MW; 003ABCC90FE732F9 CRC64;

Query Match 58.8%; Score 40; DB 2; Length 431;
Best Local Similarity 75.0%; Pred. No. 3e+02;
Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 KAAAVLLPVLLA 13
DB 227 KTAAMLLPVTLA 238
|||:|||||
|:::|||||

RESULT 41
Q7D1D9 PRELIMINARY; PRT; 435 AA.
AC Q7D1D9;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE AGR C 854p.
GN OrderedLocusNames=AGR C 854;
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
OX NCBI_TaxID=176299;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Cereon;
RX MEDLINE=21608551; PubMed=11743194; DOI=10.1126/science.1066803;
RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,
RA Quorillo B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,
RA Houmlel K., Gordon J., Vaudin M., Iartchouk O., Epp A., Liu F.,
RA Wollam C., Allinger M., Doughty D., Scott C., Lappas C., Markelz B.,
RA Flanagan C., Crowell C., Gursion J., Lomo C., Sear C., Strub G.,
RA Cielo C., Slater S.;
RT "Genome sequence of the plant pathogen and biotechnology agent
RT Agrobacterium tumefaciens C58."
RL Science 294:2323-2328(2001).

DR EMBL; AE007984; AAK86294.1; -.
DR InterPro; IPR002345; Lipocalin.
DR InterPro; IPR007016; Wzy_C.
DR Pfam; PF04932; Wzy_C; 1.
DR PROSITE; PS00213; LIPOCALIN; UNKNOWN_1.
SQ SEQUENCE 435 AA; 48138 MW; 433848D6B6B4FB4E CRC64;

Query Match 58.8%; Score 40; DB 2; Length 435;
Best Local Similarity 75.0%; Pred. No. 3e+02;
Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 KAAAVLLPVLLA 13
DB 231 KTAAMLLPVTLA 242
|||:|||||
|:::|||||

RESULT 42
Q70PA4 PRELIMINARY; PRT; 450 AA.
AC Q70PA4;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
OS Melittangium lichenicola.
OC Bacteria; Proteobacteria; Deltaproteobacteria; Myxococcales;
OC Cystobacteriaceae; Cystobacteraceae; Melittangium.
OX NCBI_TaxID=45;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22945398; PubMed=14583260; DOI=10.1016/j.chembiol.2003.09.012;
RA Weinig S., Hecht H.J., Mahmud T., Mueller R.;
RT "Melithiazol Biosynthesis. Further Insights into Myxobacterial
RT PKS/NRPS Systems and Evidence for a New Subclass of Methyl
RT Transferases."
RL Chem. Biol. 10:939-952(2003).
DR EMBL; AJ557546; CAD89767.1; -.
KW Hypothetical protein.
SQ SEQUENCE 450 AA; 49816 MW; 7243648A50D9BESC CRC64;

Query Match 58.8%; Score 40; DB 2; Length 450;
Best Local Similarity 53.3%; Pred. No. 3.1e+02;
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 KKAATAVLLPVLLAAP 15
DB 4 RRLSALLLPVLFAP 18
|||:|||||
|:::|||||

RESULT 43
Q82DE6 PRELIMINARY; PRT; 458 AA.
AC Q82DE6;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=SAV5036;
OS Streptomyces avermitilis.
OC Bacteria; Actinobacteria; Actinomycetaceae; Streptomyces.
OX NCBI_TaxID=33903;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MA-4680;
RX MEDLINE=21477403; PubMed=11572948; DOI=10.1073/pnas.211433198;
RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osone T.,
RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
RT "Genome sequence of an industrial microorganism Streptomyces
RT avermitilis: deducing the ability of producing secondary
RT metabolites."
RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).

```



```

RN  [2]
RP  SEQUENCE FROM N.A.
RX  STRAIN=MA-4680;
RA  MEDLINE=22608306; PubMed=12692562;
RA  Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
RA  Sakaki Y., Hattori M., Omura S.;
RT  "Complete genome sequences and comparative analysis of the industrial
RT  microorganism Streptomyces avermitilis.";
RL  Nat. Biotechnol. 21:526-531(2003).
DR  EMBL; AP005041; BAC72748.1; -.
DR  HSSP; Q9VGN3; 1OHT.
DR  GO; GO:0008745; F:N-acetylmuramoyl-L-alanine amidase activity; IEA.
DR  GO; GO:0005515; F:protein binding; IEA.
DR  GO; GO:0009253; F:peptidoglycan catabolism; IEA.
DR  InterPro; IPR002502; Amidase_2.
DR  Pfam; PF01510; Amidase_2; 1.
DR  SMART; SM00701; GPRP; 1.
KW  Complete proteome; Hypothetical protein.
SQ  SEQUENCE 458 AA; 47722 MW; 02F837E43702989A CRC64;

Query Match      58.8%; Score 40; DB 2; Length 458;
Best Local Similarity 61.5%; Pred. No. 3.1e+02;
Matches 8; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy  3 AAVALLPVLLAAP 15
Db  14 AAALVPLTLTAP 26

RESULT 44
Q6NF46
ID  Q6NF46 PRELIMINARY; PRT; 483 AA.
AC  Q6NF46;
DT  05-JUL-2004 (TrEMBLrel. 27, Created)
DT  05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT  05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE  Phosphate acetyltransferase (EC 2.3.1.8).
GN  Name=pta; OrderedLocusNames=DIP2055;
OS  Corynebacterium diphtheriae.
OC  Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC  Corynebacterineae; Corynebacteriaceae; Corynebacterium.
OX  NCBI_TaxID=1717;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  STRAIN=Biotype gravis / NCTC 13129;
RX  MEDLINE=22965443; PubMed=14603910; DOI=10.1093/nar/gkg874;
RA  Cerdeno-Tarraga A.-M., Efstathiou A., Dover L.G., Holden M.T.G.,
RA  Fallen M.J., Bentley S.D., Besra G.S., Churcher C.M., James K.D.,
RA  De Zeyva A., Chillingworth T., Cronin A., Dowd L., Feltwell T.,
RA  Hamlin N., Holroyd S., Jagels K., Moule S., Quail M.A.,
RA  Rabinowitsch E., Rutherford K.M., Thomson N.R., Unwin L.,
RA  Whitehead S., Barrell B.G., Parkhill J.;
RT  "The complete genome sequence and analysis of Corynebacterium
RT  diphtheriae NCTC13129.";
RL  Nucleic Acids Res. 31:6516-6523(2003).
DR  EMBL; BX248360; CAES0581.1; -.
DR  GO; GO:0008959; F:phosphate acetyltransferase activity; IEA.
DR  GO; GO:0016740; F:transferase activity; IEA.
DR  GO; GO:0008152; P:metabolism; IEA.
DR  InterPro; IPR004614; PtaPactrans.
DR  Pfam; PF01515; PTA_PTB; 1.
DR  TIGRFAMs; TIGR00651; pta; 1.
KW  Acyltransferase; Complete proteome; Transferase.
SQ  SEQUENCE 483 AA; 51065 MW; 6B43ACFE54ED9BC1 CRC64;

Query Match      58.8%; Score 40; DB 2; Length 483;
Best Local Similarity 57.1%; Pred. No. 3.3e+02;
Matches 8; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

Qy  2 KAAAVLLPVLLAAP 15
Db  ||:||| ||::||

```

```

Db  147 KASAEAFVMSAP 160

RESULT 45
MM14_PIG
ID  MM14_PIG STANDARD; PRT; 580 AA.
AC  Q9XT90;
DT  16-OCT-2001 (Rel. 40, Created)
DT  16-OCT-2001 (Rel. 40, Last sequence update)
DT  25-OCT-2004 (Rel. 45, Last annotation update)
DE  Matrix metalloproteinase-14 precursor (EC 3.4.24.80) (MMP-14)
DE  (Membrane-type matrix metalloproteinase 1) (MT-MMP 1) (MTMMP1).
GN  Name=MM14;
OS  Sus scrofa (Pig).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX  NCBI_TaxID=9823;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=9905929; PubMed=9881602; DOI=10.1016/S0945-053X(98)90098-1;
RA  Caron C., Xue J., Bartlett J.D.;
RT  "Expression and localization of membrane type 1 matrix
RT  metalloproteinase in tooth tissues.";
RL  Matrix Biol. 17:501-511(1998).
CC  -!- FUNCTION: Seems to specifically activate progelatinase A. May thus
CC  trigger invasion by tumor cells by activating progelatinase A on
CC  the tumor cell surface (By similarity). May play a role in the
CC  biomineralization of enamel and dentin.
CC  -!- CATALYTIC ACTIVITY: Endopeptidase activity. Activates
CC  progelatinase A by cleavage of the propeptide at 37-Asn-Ileu-38.
CC  Other bonds hydrolyzed include 35-Gly-Ile-36 in the propeptide
CC  of collagenase 3, and 341-Asn-Ileu-342, 441-Asp-Ileu-442 and
CC  354-Gln-Thr-355 in the aggrecan interglobular domain.
CC  -!- COFACTOR: Binds 1 zinc ion per subunit, calcium (by similarity).
CC  -!- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
CC  -!- TISSUE SPECIFICITY: Highly expressed in developing tooth tissues.
CC  -!- SIMILARITY: Belongs to the peptidase M10A family.
CC  -!- SIMILARITY: Contains 1 hemopexin-like domain.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; AF067419; AAD38324.1; -.
DR  HSPS; P50281; 1BQQ.
DR  MEROPS; M10.014; -.
DR  InterPro; IPR000585; Hemopexin.
DR  InterPro; IPR001818; Pept_M10A_M12B.
DR  InterPro; IPR006025; Pept_M_Zn_BS.
DR  InterPro; IPR006026; Peptidase_M.
DR  Pfam; PF00045; Hemopexin; 4.
DR  Pfam; PF00413; Peptidase_M10; 1.
DR  PRINTS; PR00138; Peptidase_M10_N; 1.
DR  SMART; SM00120; HX; 4.
DR  SMART; SM00235; ZnMG; 1.
DR  PROSITE; PS00546; CYSTEINE_SWITCH; 1.
DR  PROSITE; PS00024; HEMOPEXIN; 1.
DR  PROSITE; PS00142; ZINC_PROTEASE; 1.
KW  Calcium; Hydrolase; Metalloprotease; Signal; Transmembrane; Zinc;
KW  Zymogen.
FT  SIGNAL 1 28 Potential.
FT  PROPEP 29 109 Activation peptide.
FT  CHAIN 110 580 Matrix metalloproteinase-14.
FT  DOMAIN 110 539 Extracellular (Potential).
FT  TRANSMEM 540 560 Potential.
FT  DOMAIN 561 580 Cytoplasmic (Potential).
FT  DOMAIN 314 509 Hemopexin-like.

```

FT SITE 91 91 Cysteine switch (Potential).  
 FT METAL 237 237 Zinc (catalytic) (By similarity).  
 FT ACT SITE 238 238 By similarity.  
 FT METAL 241 241 Zinc (catalytic) (By similarity).  
 FT METAL 247 247 Zinc (catalytic) (By similarity).  
 FT DISULFID 317 506 By similarity.  
 SQ SEQUENCE 580 AA; 65934 MW; B7B2C2C569A96CAC CRC64;  
 Query Match 58.8%; Score 40; DB 1; Length 580;  
 Best Local Similarity 90.8%; Pred. NO. 3.9e+02;  
 Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 3 AAVALPVL 12  
 Db 537 AAVALPVL 546  
 RESULT 46  
 MM14 HUMAN STANDARD; PRT; 582 AA.  
 AC P50251; Q92678;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 25-OCT-2004 (Rel. 45, Last annotation update)  
 DE Matrix metalloproteinase-14 precursor (EC 3.4.24.80) (MMP-14)  
 DE (Membrane-type matrix metalloproteinase 1) (MT-MMP 1) (MTMMP1)  
 DE (Membrane-type-1 matrix metalloproteinase) (MT1-MMP) (MT1MMP) (MMP-X1).  
 GN Names=MMP14;  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Placenta;  
 RX MEDLINE=94286011; PubMed=8015608; DOI=10.1038/370061a0;  
 RA Sato H., Takino T., Okada Y., Cao J., Shinagawa A., Yamamoto E.,  
 Seiki M.;  
 RA "A matrix metalloproteinase expressed on the surface of invasive  
 tumour cells.";  
 RT Nature 370:61-65 (1994).  
 RL [2]  
 RN SEQUENCE FROM N.A.  
 RP TISSUE=Placenta;  
 RX MEDLINE=95237627; PubMed=7721107; DOI=10.1016/0378-1119(94)00637-8;  
 RA Takino T., Sato H., Yamamoto E., Seiki M.;  
 RT "Cloning of a human gene potentially encoding a novel matrix  
 metalloproteinase having a C-terminal transmembrane domain.";  
 RL Gene 155:293-298 (1995).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=95224014; PubMed=7708715;  
 RA Okada A., Bellocq J.P., Rouyer N., Chenard M.P., Rio M.C., Chambon P.,  
 Basset P.;  
 RT "Membrane-type matrix metalloproteinase (MT-MMP) gene is expressed in  
 stromal cells of human colon, breast, and head and neck carcinomas.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 92:2730-2734 (1995).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Lung;  
 RX MEDLINE=95377289; PubMed=7649159;  
 RA Will H., Hinzmann B.;  
 RT "cDNA sequence and mRNA tissue distribution of a novel human matrix  
 metalloproteinase with a potential transmembrane segment.";  
 RL Eur. J. Biochem. 231:602-608 (1995).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RA Luo G.X., Reisfeld R.A., Strongin A.Y.;  
 RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RA Lohi J.L., Westermarck J., Kaeheri V.M., Keeki-Oja J.;

RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.  
 RN [7]  
 RP SEQUENCE OF 112-116.  
 RX PubMed=8804434;  
 RA Sato H., Kinoshita T., Takino T., Nakayama K., Seiki M.;  
 RT "Activation of a recombinant membrane type 1-matrix metalloproteinase  
 (MT1-MMP) by furin and its interaction with tissue inhibitor of  
 metalloproteinases (TIMP)-2.";  
 RT FEBS Lett. 393:101-104 (1996).  
 RL CC  
 CC FUNCTION: Seems to specifically activate progelatinase A. May thus  
 trigger invasion by tumor cells by activating progelatinase A on  
 the tumor cell surface.  
 CC CC  
 CC CATALYTIC ACTIVITY: Endopeptidase activity. Activates  
 progelatinase A by cleavage of the propeptide at 37-Asn-Ileu-38.  
 CC Other bonds hydrolyzed include 35-Gly-Ile-36 in the propeptide  
 CC of collagenase 3, and 341-Asn-Ile-342, 441-Asp-Ileu-442 and  
 CC 354-Gln-Thr-355 in the aggrecan interglobular domain.  
 CC CC  
 CC COFACTOR: Binds 1 zinc ion per subunit, calcium (By similarity).  
 CC CC  
 CC SUBCELLULAR LOCATION: Type I membrane protein (Potential).  
 CC CC  
 CC TISSUE SPECIFICITY: In stromal cells of colon, breast, and head  
 and neck.  
 CC CC  
 CC SIMILARITY: Belongs to the peptidase M10A family.  
 CC CC  
 CC SIMILARITY: Contains 1 hemopexin-like domain.  
 CC CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC CC  
 CC EMBL; D26512; BAA05519.1; -;  
 CC EMBL; X83535; CAA58519.1; -;  
 CC EMBL; 248481; CAA88372.1; -;  
 CC EMBL; U41078; AAA83770.1; -;  
 CC EMBL; X90925; CAA62432.1; -;  
 CC PIR; I38028; I38028;  
 CC PDB; 1EQQ; X-ray; M=114-287.  
 CC PDB; 1BUV; X-ray; M=114-287.  
 CC MEROPS; M10.014; -;  
 CC Genew; HGNC:7160; MMP14.  
 CC MIM; 600754; -;  
 CC GO; GO:0005887; C:integral to plasma membrane; TAS.  
 CC GO; GO:0004222; F:metalloendopeptidase activity; TAS.  
 CC GO; GO:0008270; F:zinc ion binding; TAS.  
 CC GO; GO:0006508; P:proteolysis and peptidolysis; TAS.  
 CC InterPro; IPR000585; Hemopexin.  
 CC InterPro; IPR001818; Pept\_M10A\_M12B.  
 CC InterPro; IPR006025; Pept\_M\_Zn\_BS.  
 CC InterPro; IPR006026; Peptidase\_M.  
 CC Pfam; PF00045; Hemopexin; 4.  
 CC Pfam; PF00413; Peptidase\_M10; 1.  
 CC Pfam; PF03933; Peptidase\_M10\_N; 1.  
 CC PRINTS; PR00138; MATRXIN.  
 CC SMART; SM00120; HX; 4.  
 CC SMART; SM00235; ZnMc; 1.  
 CC PROSITE; PS00546; CYSTEINE\_SWITCH; 1.  
 CC PROSITE; PS00024; HEMOPEXIN; 1.  
 CC PROSITE; PS00142; ZINC\_PROTEASE; 1.  
 CC 3D-structure; Calcium; Direct protein sequencing; Hydrolase;  
 CC Metalloprotease; Signal; Transmembrane; Zinc; Zymogen.  
 CC SIGNAL  
 CC 1 20  
 CC PROPEP 21 111 Activation peptide.  
 CC CHAIN 112 582 Matrix metalloproteinase-14.  
 CC DOMAIN 112 541 Extracellular (Potential).  
 CC TRANSMEM 542 562 Potential.  
 CC DOMAIN 563 582 Cytoplasmic (Potential).  
 CC DOMAIN 316 511 Hemopexin-like.  
 CC SITE 93 93 Cysteine switch (Potential).  
 CC METAL 239 239 Zinc (catalytic) (By similarity).  
 CC ACT\_SITE 240 240 By similarity.  
 CC METAL 243 243 Zinc (catalytic) (By similarity).

during mouse embryogenesis.";  
 J. Biol. Chem. 272:25511-25517(1997).  
 [4]  
 SEQUENCE FROM N.A.  
 STRAIN=CD-1; TISSUE=Kidney;  
 MEDLINE=98311877; PubMed=9648071;  
 Ota K., Stetler-Stevenson W.G., Yang Q., Kumar A., Wada J.,  
 Kashiwara N., Wallner E.I., Kanwar Y.S.;  
 "Cloning of murine membrane-type-1-matrix metalloproteinase (MT-1-MMP)  
 and its metalloproteinase-type-1-matrix metalloproteinase (MT-1-MMP)  
 its inhibitor.";  
 Kidney Int. 54:131-142(1998).  
 [5]  
 FUNCTION.  
 MEDLINE=99449306; PubMed=10520996; DOI=10.1016/S0092-8674(00)80064-1;  
 Holmbeck K., Bianco P., Caterina J., Yamada S., Kromer M.,  
 Kuznetsov S.A., Mankani M., Robey P.G., Poole A.R., Fidoux I.,  
 Ward J.M., Birkedal-Hansen H.;  
 "MT1-MMP-deficient mice develop dwarfism, osteopenia, arthritis, and  
 connective tissue disease due to inadequate collagen turnover.";  
 Cell 99:81-92(1999).  
 -!- FUNCTION: Endopeptidase that degrades various components of the  
 extracellular matrix, such as collagen. Activates progelatinase A.  
 Essential for pericellular collagenolysis and modeling of skeletal  
 and extracellular connective tissues during development.  
 -!- CATALYTIC ACTIVITY: Endopeptidase activity. Activates  
 progelatinase A by cleavage of the propeptide at 37-Asn-Ileu-38.  
 Other bonds hydrolyzed include 35-Gly-Ile-36 in the propeptide  
 of collagenase 3, and 341-Asn-Phe-342, 441-Asp-Ileu-442 and  
 354-Gln-Thr-355 in the aggrecan interglobular domain.  
 -!- SUBCELLULAR LOCATION: Type I membrane protein (Potential).  
 -!- TISSUE SPECIFICITY: Highly expressed in placenta, kidney, heart,  
 lung, embryonic skeletal and pericardial tissues.  
 -!- DEVELOPMENTAL STAGE: Not detected before day 10.5. At day 12.5,  
 arteries, expressed at lower levels in the myocardium,  
 craniofacial mesenchyme, nasal epithelium and liver capsule. At  
 days 14.5 and 17.5, expressed in the musculoskeletal system, and  
 ossification areas, with continued expression in the arterial  
 tunica media.  
 -!- PTM: The precursor is cleaved by a furin endopeptidase (By  
 similarity).  
 -!- SIMILARITY: Belongs to the peptidase M10A family.  
 -!- SIMILARITY: Contains 1 hemopexin-like domain.  
 -----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 -----  
 EMBL; X835336; CAA58520.2; -;  
 EMBL; AF022432; AAB86602.1; -;  
 EMBL; AF022424; AAB86602.1; JOINED.  
 EMBL; AF022425; AAB86602.1; JOINED.  
 EMBL; AF022426; AAB86602.1; JOINED.  
 EMBL; AF022427; AAB86602.1; JOINED.  
 EMBL; AF022428; AAB86602.1; JOINED.  
 EMBL; AF022429; AAB86602.1; JOINED.  
 EMBL; AF022430; AAB86602.1; JOINED.  
 EMBL; AF022431; AAB86602.1; JOINED.  
 EMBL; U54984; AAB51753.1; -;  
 HSSP; P50281; 1BQQ.  
 MEROPS; M10.014; -;  
 MGD; MGI:101900; Mmp14.  
 InterPro; IPR000585; Hemopexin.  
 InterPro; IPR001818; Pept\_M10A\_M12B.  
 InterPro; IPR006025; Pept\_M\_Zn\_BS.  
 InterPro; IPR006026; Peptidase\_M.  
 Pfam; PF00045; Hemopexin; 4.

Zinc (catalytic) (By similarity).  
 By similarity.  
 K -> E (in Ref. 2, 4, 5 and 6).  
 S -> P (in Ref. 6).  
 249  
 319  
 338  
 500  
 500  
 123  
 125  
 127  
 128  
 133  
 135  
 137  
 154  
 158  
 160  
 169  
 170  
 178  
 182  
 198  
 202  
 208  
 211  
 216  
 218  
 219  
 222  
 225  
 225  
 232  
 232  
 233  
 243  
 244  
 246  
 253  
 254  
 260  
 260  
 267  
 273  
 276  
 277  
 278  
 282  
 283  
 582 AA; EFCEDE6A41116F9 CRG64;  
 58.8%; Score 40; DB 1; Length 582;  
 Best Local Similarity 90.0%; Pred. No. 3.9e+02;  
 Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 3 AAVALLPVLL 12  
 539 AAVALLPVLL 548  
 Query Match 58.8%; Score 40; DB 1; Length 582;  
 Best Local Similarity 90.0%; Pred. No. 3.9e+02;  
 Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 3 AAVALLPVLL 12  
 539 AAVALLPVLL 548  
 RESULT 47  
 MM14\_MOUSE  
 ID MM14\_MOUSE STANDARD; PRT; 582 AA.  
 AC P53690; O08645; O35369;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 25-OCT-2004 (Rel. 45, Last annotation update)  
 DE Matrix metalloproteinase-14 precursor (EC 3.4.24.80) (MMP-14)  
 DE (Membrane-type matrix metalloproteinase 1) (MT-MMP 1) (MTMMP1)  
 DE (Membrane-type-1 matrix metalloproteinase) (MT1-MMP) (MTMMP) (MMP-X1)  
 DE (MMP)  
 GN Name=Mmp14; Synonyms=Mtmmp;  
 OS Mus musculus (Mouse);  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 [1]  
 RN SEQUENCE FROM N.A.  
 RP MEDLINE=95224014; PubMed=7708715;  
 RA Okada A., Bellocq J.P., Rouyer N., Chenard M.P., Rio M.C., Chambon P.,  
 Basset P.;  
 RT "Membrane-type matrix metalloproteinase (MT-MMP) gene is expressed in  
 stromal cells of human colon, breast, and head and neck carcinomas.";  
 Proc. Natl. Acad. Sci. U.S.A. 92:2730-2734(1995).  
 [2]  
 RN REVISIONS.  
 RP Odaka A.;  
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.  
 [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=129/Sv;  
 RX MEDLINE=97467335; PubMed=9325265; DOI=10.1074/jbc.272.41.25511;  
 RA Apte S.S., Fukui N., Beier D.R., Olsen B.R.;  
 RT "The matrix metalloproteinase-14 (MMP-14) gene is structurally  
 distinct from other MMP genes and is co-expressed with the TIMP-2 gene

```

DR Pfam; PF00413; Peptidase M10; 1.
DR Pfam; PF03933; Peptidase_M10_N; 1.
DR PRINTS; PR00138; MATRXIN.
DR SMART; SM00120; HX; 4.
DR SMART; SM00235; ZnMc; 1.
DR PROSITE; PS00546; CYSTEINE_SWITCH; 1.
DR PROSITE; PS00024; HEMOPEXIN; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
KW Calcium; Collagen degradation; Extracellular matrix; Hydrolase;
KW Metalloprotease; Signal; Transmembrane; Zinc; Zymogen.
FT SIGNAL 1 20 Potential.
FT PROPEP 21 111 By similarity.
FT CHAIN 112 582 Matrix metalloproteinase-14.
FT DOMAIN 112 541 Extracellular (Potential).
FT TRANSMEM 542 562 Potential.
FT DOMAIN 563 582 Cytoplasmic (Potential).
FT DOMAIN 316 511 Hemopexin-like.
FT SITE 93 93 Cysteine switch (Potential).
FT METAL 239 239 Zinc (catalytic) (By similarity).
FT ACT_SITE 240 240 By similarity.
FT METAL 243 243 Zinc (catalytic) (By similarity).
FT METAL 249 249 Zinc (catalytic) (By similarity).
FT DISULFID 319 508 By similarity.
FT CONFLICT 29 29 Q -> K (in Ref. 2).
FT CONFLICT 268 268 K -> N (in Ref. 2).
FT CONFLICT 270 270 L -> V (in Ref. 2).
FT CONFLICT 275 275 E -> D (in Ref. 2).
FT CONFLICT 292 296 RCLLN -> KMPPP (in Ref. 2).
FT CONFLICT 298 300 GQP -> RTT (in Ref. 2).
FT CONFLICT 302 308 GULFRIS -> RTFIPDK (in Ref. 2).
FT CONFLICT 310 310 G -> R (in Ref. 2).
FT CONFLICT 317 317 K -> N (in Ref. 2).
FT CONFLICT 329 329 F -> L (in Ref. 2).
FT CONFLICT 360 360 L -> P (in Ref. 2).
SQ SEQUENCE 582 AA; 65935 MW; 3AB355158D4DD175 CRC64;

Query Match 58.8%; Score 40; DB 1; Length 582;
Best Local Similarity 90.0%; Pred. No. 3.9e+02;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 AAAYLLPVLL 12
|||||
Db 539 AAAYLLPVLL 548

RESULT 48
MM14 RABIT STANDARD; PRT; 582 AA.
AC Q95270; P79225;
DT 01-NOV-1997 (Rel. 35, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DE Matrix metalloproteinase-14 precursor (EC 3.4.24.80) (MMP-14)
DE (Membrane-type matrix metalloproteinase 1) (MT-MMP 1) (MTMMP1).
DE (Membrane-type-1 matrix metalloproteinase) (MT1-MMP) (MT1MMP).
GN Name=MP14;
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]_TaxID=9986;
RP SEQUENCE FROM N.A.
RC STRAIN=New Zealand white; TISSUE=Vascular smooth muscle;
RA Wang H.; Keiser J.;
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 1-572 FROM N.A.
RC STRAIN=New Zealand white;
RA Sato T.;
RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Seems to specifically activate progelatinase A. May thus
trigger invasion by tumor cells by activating progelatinase A on

```

```

CC the tumor cell surface (By similarity).
CC -!- CATALYTIC ACTIVITY: Endopeptidase activity. Activates
progelatinase A by cleavage of the propeptide at 37-Asn-Leu-38.
CC Other bonds hydrolyzed include 35-Gly-Ile-36 in the propeptide
of collagenase 3, and 341-Asn-Phe-342, 441-Asp-Leu-442 and
354-Gln-Thr-355 in the aggrecan interglobular domain.
CC -!- COFACTOR: Binds 1 zinc ion per subunit, calcium (By similarity).
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
CC -!- SIMILARITY: Belongs to the peptidase M10A family.
CC -!- SIMILARITY: Contains 1 hemopexin-like domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U83918; AA841500.1; -.
CC EMBL; U73940; AA013803.1; -.
CC HSPSP; P50281; IBQQ.
CC MEROPS; M10.014; -.
CC InterPro; IPR000585; Hemopexin.
CC InterPro; IPR001818; Pept_M10A_M12B.
CC InterPro; IPR006025; Pept_M_Zn_BS.
CC InterPro; IPR006026; Peptidase_M.
CC Pfam; PF00045; Hemopexin; 4.
CC Pfam; PF00413; Peptidase M10; 1.
CC PRINTS; PR00138; MATRXIN.
CC SMART; SM00120; HX; 4.
CC SMART; SM00235; ZnMc; 1.
CC PROSITE; PS00546; CYSTEINE_SWITCH; 1.
CC PROSITE; PS00024; HEMOPEXIN; 1.
CC PROSITE; PS00142; ZINC_PROTEASE; 1.
KW Calcium; Hydrolase; Metalloprotease; Signal; Transmembrane; Zinc;
Zymogen.
FT SIGNAL 1 20 Potential.
FT PROPEP 21 111 Activation peptide.
FT CHAIN 112 582 Matrix metalloproteinase-14.
FT DOMAIN 112 541 Extracellular (Potential).
FT TRANSMEM 542 562 Potential.
FT DOMAIN 563 582 Cytoplasmic (Potential).
FT DOMAIN 316 511 Hemopexin-like.
FT SITE 93 93 Cysteine switch (Potential).
FT METAL 239 239 Zinc (catalytic) (By similarity).
FT ACT_SITE 240 240 By similarity.
FT METAL 243 243 Zinc (catalytic) (By similarity).
FT METAL 249 249 Zinc (catalytic) (By similarity).
FT DISULFID 319 508 By similarity.
FT CONFLICT 29 29 Q -> K (in Ref. 2).
FT CONFLICT 268 268 K -> N (in Ref. 2).
FT CONFLICT 270 270 L -> V (in Ref. 2).
FT CONFLICT 275 275 E -> D (in Ref. 2).
FT CONFLICT 292 296 RCLLN -> KMPPP (in Ref. 2).
FT CONFLICT 298 300 GQP -> RTT (in Ref. 2).
FT CONFLICT 302 308 GULFRIS -> RTFIPDK (in Ref. 2).
FT CONFLICT 310 310 G -> R (in Ref. 2).
FT CONFLICT 317 317 K -> N (in Ref. 2).
FT CONFLICT 329 329 F -> L (in Ref. 2).
FT CONFLICT 360 360 L -> P (in Ref. 2).
SQ SEQUENCE 582 AA; 65963 MW; 844624B0AF1B6812 CRC64;

Query Match 58.8%; Score 40; DB 1; Length 582;
Best Local Similarity 90.0%; Pred. No. 3.9e+02;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 3 AAAYLLPVLL 12
|||||
Db 539 AAAYLLPVLL 548

```

```

RESULT 49
MM14 RAT
ID MM14 RAT STANDARD; PRT; 582 AA.
AC Q10739;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Matrix metalloproteinase-14 precursor (EC 3.4.24.80) (MMP-14)
DE (Membrane-type matrix metalloproteinase 1) (MT-MMP 1) (MTMMP1)
DE (Membrane-type-1 matrix metalloproteinase) (MT1-MMP) (MTMMP) (MT-MMP)
DE MMP).
CN Name=Mmp14; Synonyms=Mttmp;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95224014; PubMed=7708715;
RA Okada A., Bellocq J.P., Rouyer N., Chenard M.P., Rio M.C., Chambon P.,
RA Basset P.;
RT "Membrane-type matrix metalloproteinase (MT-MMP) gene is expressed in
RT stromal cells of human colon, breast, and head and neck carcinomas.";
RL Proc. Natl. Acad. Sci. U.S.A. 92:2730-2734(1995).
RN [2]
RP SEQUENCE FROM N.A.
RA Cossins J., Clements J., Catlin G., Wells G.;
RL Submitted (SEP-1995) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Seems to specifically activate progelatinase A. May thus
CC trigger invasion by tumor cells by activating progelatinase A on
CC the tumor cell surface.
CC -!- CATALYTIC ACTIVITY: Endopeptidase activity. Activates
CC progelatinase A by cleavage of the propeptide at 37-Asn-|-Leu-38.
CC Other bonds hydrolyzed include 35-Gly-|-Ile-36 in the propeptide
CC of collagenase 3, and 341-Asn-|-Phe-342, 441-Asp-|-Leu-442 and
CC 354-Gln-|-Thr-355 in the aggrecan interglobular domain.
CC -!- COPACITOR: Binds 1 zinc ion per subunit, calcium (By similarity).
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
CC -!- SIMILARITY: Belongs to the peptidase M10A family.
CC -!- SIMILARITY: Contains 1 hemopexin-like domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; X83537; CAA58521.1; -.
CC EMBL; X91785; CAA62897.1; -.
CC PIR; I84471; I84471.
CC HSSP; P50281; 1BQQ.
CC MEROPS; M10.014; -.
CC RGD; 620198; Mmp14.
CC InterPro; IPR005585; Hemopexin.
CC InterPro; IPR001818; Pept_M10A_M12B.
CC InterPro; IPR006025; Pept_M_Zn_BS.
CC InterPro; IPR006026; Peptidase_M.
CC Pfam; PF00045; Hemopexin; 4.
CC Pfam; PF00413; Peptidase_M10; 1.
CC Pfam; PF03933; Peptidase_M10_N; 1.
CC PRINTS; PR00138; MATRIXIN.
CC SMART; SM00120; HX; 4.
CC SMART; SM00235; ZnMc; 1.
CC PROSITE; PS00546; CYSTEINE_SWITCH; 1.
CC PROSITE; PS00024; HEMOPEXIN; 1.
CC PROSITE; PS00142; ZINC_PROTEASE; 1.
CC KX Calcium; Hydrolase; Metalloprotease; Signal; Transmembrane; Zinc;
CC Zymogen.
CC
CC SIGNAL
CC 1 20 Potential.
CC 21 111 Activation peptide.
CC PROPEP 112 582 Matrix metalloproteinase-14.
CC CHAIN

```

```

FT DOMAIN 112 541 Extracellular (Potential).
FT TRANSMEM 542 562 Potential.
FT DOMAIN 563 582 Cytoplasmic (Potential).
FT DOMAIN 316 511 Hemopexin-like.
FT SITE 93 93 Cysteine switch (Potential).
FT METAL 239 239 Zinc (catalytic) (By similarity).
FT ACT_SITE 240 240 By similarity.
FT METAL 243 243 Zinc (catalytic) (By similarity).
FT METAL 249 249 Zinc (catalytic) (By similarity).
FT DISULFID 319 508 By similarity.
FT CONFLICT 68 68 I -> M (in Ref. 2).
FT CONFLICT 255 255 D -> A (in Ref. 2).
SQ SEQUENCE 582 AA; 66106 MW; 8B40FDD9999CA80C CRC64;

Query Match 58.8%; Score 40; DB 1; Length 582;
Best Local Similarity 90.0%; Pred. No. 3.9e+02;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 3 AAAYLLPVLL 12
Db 539 AAAYLLPVLL 548
|||||

RESULT 50
Q9GLE4 PRELIMINARY; PRT; 582 AA.
AC Q9GLE4;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Membrane-type matrix metalloprotease 1.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RA Zhang B., Yan L., Moses M.A., Fang J., Miao H., Teang P.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF290429; AAG28170.1; -.
DR HSSP; P50281; 1BQQ.
DR MEROPS; M10.014; -.
DR GO; GO:0005578; C:extracellular matrix (sensu Metazoa); IEA.
DR GO; GO:0004222; F:metalloendopeptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR000585; Hemopexin.
DR InterPro; IPR006026; Peptidase_M.
DR InterPro; IPR001818; Pept_M10A_M12B.
DR InterPro; IPR006025; Pept_M_Zn_BS.
DR InterPro; IPR009070; PGSD-like.
DR Pfam; PF00045; Hemopexin; 4.
DR Pfam; PF03933; Peptidase_M10_N; 1.
DR PRINTS; PR00138; MATRIXIN.
DR SMART; SM00120; HX; 4.
DR SMART; SM00235; ZnMc; 1.
DR PROSITE; PS00546; CYSTEINE_SWITCH; 1.
DR PROSITE; PS00024; HEMOPEXIN; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.
KW Metalloprotease; Protease.
SQ SEQUENCE 582 AA; 65882 MW; 65174CE65D4040E1 CRC64;

Query Match 58.8%; Score 40; DB 2; Length 582;
Best Local Similarity 90.0%; Pred. No. 3.9e+02;
Matches 9; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 3 AAAYLLPVLL 12
Db 539 AAAYLLPVLL 548
|||||

Search completed: June 2, 2005, 01:39:49
Job time : 129.406 secs

```



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 2, 2005, 01:21:36 ; Search time 109.688 Seconds  
(without alignments)  
52.890 Million cell updates/sec

Title: US-10-764-235-2

Perfect score: 68

Sequence: 1 KKAARVLLPVLAAAP 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 750 summaries

Database :

A\_Geneseq\_16Dec04:\*

1: Geneseqp1980s:\*

2: Geneseqp1990s:\*

3: Geneseqp2000s:\*

4: Geneseqp2001s:\*

5: Geneseqp2002s:\*

6: Geneseqp2003as:\*

7: Geneseqp2003bs:\*

8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	68	100.0	15	6	ABP70234 Amino aci
2	68	100.0	15	6	ABP70214 Amino aci
3	68	100.0	15	8	ABG75426 Membrane
4	68	100.0	16	4	AAU00633 Human mem
5	68	100.0	16	6	ABP70215 Amino aci
6	68	100.0	16	7	ADB88776 Membrane
7	68	100.0	19	4	AAU00655 Human tar
8	68	100.0	19	4	AAU00634 Human mem
9	68	100.0	19	7	ADB88777 Membrane
10	68	100.0	19	8	ABG75428 Membrane
11	68	100.0	20	6	ABP70216 Amino aci
12	68	100.0	21	4	AAU00632 Human mem
13	68	100.0	21	7	ADB88775 Membrane
14	64	94.1	16	4	AAU00635 Human mem
15	64	94.1	16	7	ADB88778 Membrane
16	64	94.1	16	8	ABG75429 Membrane
17	63	92.6	14	8	ABG75427 Membrane
18	63	92.6	15	6	ABP70217 Amino aci
19	61	89.7	15	6	ABP70233 Amino aci
20	58	85.3	21	8	ADK50911 Human C35
21	58	85.3	21	8	ADK50919 Human C35
22	58	85.3	21	8	ADK50915 Human C35
23	58	85.3	33	8	ADK50916 Human C35
24	58	85.3	33	8	ADK50912 Human C35
25	58	85.3	33	8	ADK50920 Human C35

26	57	83.8	18	8	ADR40510	Adr40510 Peptide c
27	57	83.8	18	8	ADR40509	Adr40509 Peptide c
28	57	83.8	21	8	ADK50831	Human C35
29	57	83.8	33	8	ADK50832	Human C35
30	56	82.4	21	8	ADK50791	Human C35
31	56	82.4	33	8	ADK50792	Human C35
32	55	80.9	17	8	ADK50791	Human C35
33	55	80.9	17	8	ADK50792	Human C35
34	54	79.4	12	2	AAU00644	Human mem
35	54	79.4	12	2	AAU00644	Human mem
36	54	79.4	12	4	AAU00644	Human mem
37	54	79.4	12	5	ABG78983	Cell pene
38	54	79.4	12	5	AAU78350	Grb2 SH2
39	54	79.4	12	5	ABG78983	Cell pene
40	54	79.4	12	5	ABG78983	Cell pene
41	54	79.4	12	5	ABG78983	Cell pene
42	54	79.4	12	5	ABG78983	Cell pene
43	54	79.4	12	5	ABG78983	Cell pene
44	54	79.4	12	5	ABG78983	Cell pene
45	54	79.4	12	5	ABG78983	Cell pene
46	54	79.4	12	5	ABG78983	Cell pene
47	54	79.4	12	5	ABG78983	Cell pene
48	54	79.4	12	5	ABG78983	Cell pene
49	54	79.4	12	5	ABG78983	Cell pene
50	54	79.4	12	5	ABG78983	Cell pene
51	54	79.4	12	5	ABG78983	Cell pene
52	54	79.4	12	5	ABG78983	Cell pene
53	54	79.4	12	5	ABG78983	Cell pene
54	54	79.4	12	5	ABG78983	Cell pene
55	54	79.4	12	5	ABG78983	Cell pene
56	54	79.4	12	5	ABG78983	Cell pene
57	54	79.4	12	5	ABG78983	Cell pene
58	54	79.4	12	5	ABG78983	Cell pene
59	54	79.4	12	5	ABG78983	Cell pene
60	54	79.4	12	5	ABG78983	Cell pene
61	54	79.4	12	5	ABG78983	Cell pene
62	54	79.4	12	5	ABG78983	Cell pene
63	54	79.4	12	5	ABG78983	Cell pene
64	54	79.4	12	5	ABG78983	Cell pene
65	54	79.4	12	5	ABG78983	Cell pene
66	54	79.4	12	5	ABG78983	Cell pene
67	54	79.4	12	5	ABG78983	Cell pene
68	54	79.4	12	5	ABG78983	Cell pene
69	54	79.4	12	5	ABG78983	Cell pene
70	54	79.4	12	5	ABG78983	Cell pene
71	54	79.4	12	5	ABG78983	Cell pene
72	54	79.4	12	5	ABG78983	Cell pene
73	54	79.4	12	5	ABG78983	Cell pene
74	54	79.4	12	5	ABG78983	Cell pene
75	54	79.4	12	5	ABG78983	Cell pene
76	54	79.4	12	5	ABG78983	Cell pene
77	54	79.4	12	5	ABG78983	Cell pene
78	54	79.4	12	5	ABG78983	Cell pene
79	54	79.4	12	5	ABG78983	Cell pene
80	54	79.4	12	5	ABG78983	Cell pene
81	54	79.4	12	5	ABG78983	Cell pene
82	54	79.4	12	5	ABG78983	Cell pene
83	54	79.4	12	5	ABG78983	Cell pene
84	54	79.4	12	5	ABG78983	Cell pene
85	54	79.4	12	5	ABG78983	Cell pene
86	54	79.4	12	5	ABG78983	Cell pene
87	54	79.4	12	5	ABG78983	Cell pene
88	54	79.4	12	5	ABG78983	Cell pene
89	54	79.4	12	5	ABG78983	Cell pene
90	54	79.4	12	5	ABG78983	Cell pene
91	54	79.4	12	5	ABG78983	Cell pene
92	54	79.4	12	5	ABG78983	Cell pene
93	54	79.4	12	5	ABG78983	Cell pene
94	54	79.4	12	5	ABG78983	Cell pene
95	54	79.4	12	5	ABG78983	Cell pene
96	54	79.4	12	5	ABG78983	Cell pene
97	54	79.4	12	5	ABG78983	Cell pene
98	54	79.4	12	5	ABG78983	Cell pene

99	54	79.4	33	8	ADK50796	Human C35	172	41	60.3	17	4	AAE12503	Membrane
100	54	79.4	33	8	ADK50888	Human C35	173	41	60.3	19	3	AAy67956	Karposi s
101	54	79.4	33	8	ADK50800	Human C35	174	41	60.3	20	3	AAy67963	Karposi s
102	54	79.4	33	8	ADK50812	Human C35	175	41	60.3	20	4	AAE12478	Membrane
103	54	79.4	33	8	ADK50816	Human C35	176	41	60.3	20	4	AAE12479	Membrane
104	54	79.4	33	8	ADK50824	Human C35	177	41	60.3	20	4	AAE12477	Membrane
105	54	79.4	34	8	ADK50924	Human C35	178	41	60.3	22	3	AAy67959	Karposi s
106	54	79.4	34	8	ADK50908	Human C35	179	41	60.3	22	3	AAy67958	Karposi s
107	54	79.4	34	8	ADK50840	Human C35	180	41	60.3	53	4	AAU55087	Propionib
108	54	79.4	37	8	ADK50848	Human C35	181	41	60.3	53	6	ABM51606	Propionib
109	54	79.4	38	8	ADK50884	Human C35	182	41	60.3	578	7	ABO69542	Pseudomon
110	54	79.4	38	8	ADK50855	Human C35	183	40	58.8	30	2	AAr75651	Human pla
111	54	79.4	39	8	ADK50880	Human C35	184	40	58.8	108	4	AAU67736	Propionib
112	54	79.4	39	8	ADK50828	Human C35	185	40	58.8	108	6	ABM64255	Propionib
113	54	79.4	42	8	ADK50859	Human C35	186	40	58.8	193	8	ADN47148	thermococ
114	54	79.4	45	8	ADK50852	Human C35	187	40	58.8	455	6	ABU26112	Protein e
115	54	79.4	49	8	ADK50863	Human C35	188	40	58.8	456	7	ABO71756	Pseudomon
116	54	79.4	50	8	ADK50856	Human C35	189	40	58.8	549	8	ADP29329	Human sec
117	54	79.4	54	8	ADK50860	Human C35	190	40	58.8	579	2	AAr86406	Human mat
118	54	79.4	54	8	ADK50867	Human C35	191	40	58.8	582	2	AAr86407	Human mat
119	54	79.4	61	8	ADK50871	Human C35	192	40	58.8	582	2	AAr75648	Human pla
120	54	79.4	61	8	ADK50864	Human C35	193	40	58.8	582	2	AAW52134	Rabbit me
121	54	79.4	66	8	ADK50868	Human C35	194	40	58.8	582	4	AAr84616	Amino aci
122	54	79.4	67	8	ADK50875	Human C35	195	40	58.8	582	4	AAE10423	Human mat
123	54	79.4	73	8	ADK50872	Human C35	196	40	58.8	582	5	AAU84294	Human end
124	54	79.4	79	8	ADK50876	Human C35	197	40	58.8	582	5	AAE21037	Human mem
125	54	79.4	386	5	AAE22205	Human C35	198	40	58.8	582	5	AAW50865	Matrix me
126	53	77.9	14	4	AAU00637	Human mem	199	40	58.8	582	7	ADC15498	Human bas
127	53	77.9	14	7	ADB88780	Membrane	200	40	58.8	582	7	ADP33426	PRO poly
128	53	77.9	14	8	ABG75431	Membrane	201	40	58.8	582	7	ADP33426	PRO poly
129	51	75.0	17	4	AAU00654	Human mem	202	40	58.8	582	7	ADP13708	Tumor-Aes
130	51	75.0	17	7	ADB88797	Membrane	203	40	58.8	582	7	ADN95889	Human BEC
131	51	75.0	17	8	ABG75448	Membrane	204	40	58.8	582	8	ADN07703	Human mat
132	50	73.5	11	2	AAy44168	Truncated	205	40	58.8	582	8	ABM81541	Tumour-ab
133	50	73.5	11	4	AAU00649	Human mem	206	40	58.8	582	8	ADP23426	PRO poly
134	50	73.5	11	7	ADB88792	Membrane	207	40	58.8	1246	5	AAU84295	Human end
135	50	73.5	11	8	ABG75443	Membrane	208	40	58.8	1247	5	ABN90742	Human Tum
136	50	73.5	13	4	AAU00640	Human mem	209	40	58.8	1247	5	AAO17366	Human nid
137	50	73.5	13	7	ADB88783	Membrane	210	40	58.8	1247	6	ABU54449	Human tum
138	50	73.5	13	8	ABG75434	Membrane	211	40	58.8	1247	6	ADN95448	Human BEC
139	47	69.1	11	2	AAy44164	Truncated	212	40	58.8	1247	8	ADN94932	Human nid
140	47	69.1	11	4	AAU00645	Human mem	213	39	57.4	9	2	AAy44162	Truncated
141	47	69.1	11	7	ADB88788	Membrane	214	39	57.4	9	4	AAU00647	Human mem
142	47	69.1	11	8	ABG75439	Membrane	215	39	57.4	9	7	ADB88790	Membrane
143	47	69.1	12	4	AAU00646	Human mem	216	39	57.4	9	8	ABG75441	Membrane
144	47	69.1	12	7	ADB88789	Membrane	217	39	57.4	11	4	AAU00638	Human mem
145	47	69.1	12	8	ABG75440	Membrane	218	39	57.4	11	7	ADB88781	Membrane
146	47	69.1	13	4	AAU00636	Human mem	219	39	57.4	11	8	ABG75432	Membrane
147	47	69.1	13	7	ADB88779	Membrane	220	39	57.4	26	8	ADJ36102	Self-coal
148	47	69.1	13	8	ABG75430	Membrane	221	39	57.4	88	3	AAg27031	Zea nays
149	47	69.1	17	4	AAU00653	Human mem	222	39	57.4	91	3	AAg27030	Zea nays
150	47	69.1	17	7	ADB88796	Membrane	223	39	57.4	180	5	ABG66679	Human nov
151	47	69.1	17	8	ABG75447	Membrane	224	39	57.4	227	6	ABP81748	Human G p
152	46	67.6	10	2	AAy44167	Truncated	225	39	57.4	240	5	AAE16174	Human G-p
153	46	67.6	10	4	AAU00650	Human mem	226	39	57.4	250	8	ADP50658	Murine DN
154	46	67.6	10	7	ADB88793	Membrane	227	39	57.4	254	5	ABM05672	Human amy
155	46	67.6	10	8	ABG75444	Membrane	228	39	57.4	254	7	ADJ69690	Human hea
156	46	67.6	11	8	ADB88796	Membrane	229	39	57.4	266	4	AAU42505	Propionib
157	46	67.6	12	4	AAU00641	Human mem	230	39	57.4	266	6	ABM39024	Propionib
158	46	67.6	12	7	ADB88784	Membrane	231	39	57.4	273	6	ABU19701	Protein e
159	46	67.6	12	8	ABG75435	Membrane	232	39	57.4	273	6	ABU22263	Protein e
160	45	66.2	330	2	ADA48034	Rice prot	233	39	57.4	308	7	ADA49423	Multi-epi
161	43	63.2	10	2	AAy44163	Truncated	234	39	57.4	308	8	ADO24101	Epigene c
162	43	63.2	186	6	ABU41298	Protein e	235	39	57.4	405	8	ADI42391	Plant tra
163	43	63.2	191	7	ADCF0763	Bacterial	236	39	57.4	510	7	ABO68625	Pseudomon
164	43	63.2	425	7	ADG12732	Human GPC	237	39	57.4	513	4	AAg82930	S. epider
165	42	61.8	9	2	AAy44166	Truncated	238	39	57.4	529	6	ABU42668	Protein e
166	42	61.8	9	4	AAU00651	Human mem	239	39	57.4	534	6	ABU42485	Protein e
167	42	61.8	9	7	ADB88794	Membrane	240	39	57.4	534	6	ABM72922	Staphyloc
168	42	61.8	11	8	ABG75445	Membrane	241	39	57.4	551	5	ABP40156	Staphyloc
169	42	61.8	11	4	AAU00642	Human mem	242	39	57.4	551	8	ADP07202	Staphyloc
170	42	61.8	11	7	ADB88785	Membrane	243	39	57.4	560	8	ADR10453	Human pro
171	42	61.8	11	8	ABG75436	Membrane	244	39	57.4	630	8	ADO28948	Mouse nov



245	39	57.4	633	5	AAE18653	Aae18653 Human G-p	318	37	54.4	174	8	ADL76834	Adl76834 Albumin f
246	39	57.4	706	8	ADO28946	Ado28946 Human nov	319	37	54.4	198	4	AAb74770	Aab74770 Human sec
247	39	57.4	718	6	AAU54939	Aau54939 Propionib	320	37	54.4	210	8	ADO57658	Ado57658 Actinobac
248	39	57.4	718	6	ABM51458	Abm51458 Propionib	321	37	54.4	234	5	AG622209	Ag622209 Human gen
249	39	57.4	755	8	ADP30064	Adp30064 Human sec	322	37	54.4	234	5	ABG63571	Abg63571 Human alb
250	39	57.4	854	7	ABO70102	Ab070102 Pseudomon	323	37	54.4	234	8	ADL76836	Adl76836 Albumin f
251	39	57.4	874	3	AB241199	Ab241199 Human GTP	324	37	54.4	248	5	ABB06146	Abb06146 Human NS
252	39	57.4	882	7	ADB65389	Adb65389 Human pro	325	37	54.4	298	2	AAW22111	Aaw22111 P. tetrag
253	39	57.4	1387	5	AAU81520	Aau81520 Human int	326	37	54.4	321	8	ADI37028	Adi37028 Novel hum
254	39	57.4	1418	7	ABO83621	Ab083621 Pseudomon	327	37	54.4	321	8	ADL91805	Adl91805 Human PRO
255	39.5	56.6	335	4	ABB62722	Abb62722 Drosophil	328	37	54.4	321	8	ADO20061	Ado20061 Human PRO
256	38	55.9	8	2	AA44165	Aay44165 Truncated	329	37	54.4	331	7	ADF41969	Adf41969 Xanthobac
257	38	55.9	8	2	AAU00652	Aau00652 Human mem	330	37	54.4	331	2	RAY05283	Ray05283 EGF-like
258	38	55.9	8	7	ADB88795	Adb88795 Membrane	331	37	54.4	353	2	AA408064	Aay08064 Human EGF
259	38	55.9	8	7	ABG75446	Abg75446 Membrane	332	37	54.4	353	2	AA413344	Aay13344 Amino aci
260	38	55.9	10	4	AAU00643	Aau00643 Human mem	333	37	54.4	353	3	AA491870	Aay91870 Human apo
261	38	55.9	10	7	ADB88786	Adb88786 Membrane	334	37	54.4	353	3	AA483224	Aay83224 PRO211 Po
262	38	55.9	10	8	ABG75437	Abg75437 Membrane	335	37	54.4	353	3	AA488571	Aay88571 Human PRO
263	38	55.9	20	4	AAE12493	Aae12493 Membrane	336	37	54.4	353	3	RAA00169	Raa00169 PRO211 po
264	38	55.9	33	7	ADF73134	Adf73134 Transit p	337	37	54.4	353	3	ADC78322	Adc78322 Human PRO
265	38	55.9	144	5	ABB53105	Abb53105 Human ORF	338	37	54.4	353	4	AA80212	Aab80212 Human PRO
266	38	55.9	144	5	ABB53189	Abb53189 Human ORF	339	37	54.4	353	4	AA61231	Aab61231 Human TAN
267	38	55.9	169	3	AA44505	Aab44505 Human ORF	340	37	54.4	353	4	AAU12319	Aau12319 Human PRO
268	38	55.9	182	8	ADS43627	AdS43627 Bacterial	341	37	54.4	353	4	AA853075	Aab53075 Human ang
269	38	55.9	186	3	AA24009	Aeg24009 Arabidops	342	37	54.4	353	4	AA868596	Aab68596 PRO211. 4
270	38	55.9	189	3	AA24008	Aeg24008 Arabidops	343	37	54.4	353	6	ABU71590	Abu71590 Human PRO
271	38	55.9	248	7	ADB15988	Adb15988 G-coupled	344	37	54.4	353	6	ABU17763	Abu17763 Novel hum
272	38	55.9	248	8	ADL93927	Adl93927 Human G-c	345	37	54.4	353	6	ABU71445	Abu71445 Human PRO
273	38	55.9	296	8	ADL05170	Adl05170 M. catarr	346	37	54.4	353	6	ABU81017	Abu81017 Human PRO
274	38	55.9	297	7	ADF73135	Adf73135 Transit p	347	37	54.4	353	6	ABU71891	Abu71891 Human sec
275	38	55.9	303	6	ABU35394	Abu35394 Protein e	348	37	54.4	353	6	ABO01774	Ab001774 Novel hum
276	38	55.9	313	4	AAU48672	Aau48672 Propionib	349	37	54.4	353	6	ABU66717	Abu66717 Human PRO
277	38	55.9	313	6	ABM45191	Abm45191 Propionib	350	37	54.4	353	6	ABU54347	Abu54347 Human sec
278	38	55.9	323	2	AA434056	Aay34056 E. coli o	351	37	54.4	353	6	ABO32668	Ab032668 Secreted
279	38	55.9	323	2	AA434056	Aay34056 E. coli o	352	37	54.4	353	6	ABO47362	Ab047362 Human sec
280	38	55.9	340	6	ADA48210	Ada48210 Rice prot	353	37	54.4	353	6	ABU59798	Abu59798 Novel sec
281	38	55.9	340	6	ADA48610	Ada48610 Rice prot	354	37	54.4	353	6	ABO24988	Ab024988 Human sec
282	38	55.9	340	7	ADJ11590	Adj11590 Rice prot	355	37	54.4	353	6	ABU64499	Abu64499 Human sec
283	38	55.9	367	5	ABB08211	Abb08211 Escherich	356	37	54.4	353	6	ABU67345	Abu67345 Human sec
284	38	55.9	377	2	AA434057	Aay34057 E. coli o	357	37	54.4	353	6	ABO14865	Ab014865 Human sec
285	38	55.9	377	2	AA434057	Aay34057 E. coli o	358	37	54.4	353	6	ABU66993	Abu66993 Human sec
286	38	55.9	377	4	AA434057	Aay34057 E. coli o	359	37	54.4	353	6	ABU69622	Abu69622 Novel hum
287	38	55.9	377	6	ABU14744	Abu14744 Protein e	360	37	54.4	353	6	ABO14804	Ab014804 Human sec
288	38	55.9	377	8	ADN18164	Adn18164 Bacterial	361	37	54.4	353	6	ADA45815	Ada45815 Novel hum
289	38	55.9	532	6	ABU43564	Abu43564 Protein e	362	37	54.4	353	6	ADA76246	Ada76246 Human PRO
290	38	55.9	580	4	ABU52653	Abu52653 Human nuc	363	37	54.4	353	6	ADB22207	Adb22207 Human sec
291	38	55.9	580	7	ADU69927	AdJ69927 Human hea	364	37	54.4	353	6	ADA18896	Ada18896 Human PRO
292	38	55.9	592	7	ABO74092	Ab074092 Pseudomon	365	37	54.4	353	6	ADA61519	Ada61519 Homo sapi
293	38	55.9	619	4	AA664953	Aag664953 Human ATP	366	37	54.4	353	6	ADBL19304	Adb19304 Novel hum
294	38	55.9	626	4	AAU14731	Aau14731 Novel bon	367	37	54.4	353	6	ADB27845	Adb27845 Human PRO
295	38	55.9	2802	5	ADI16953	Adi16953 Chicken N	368	37	54.4	353	6	ADA86324	Ada86324 Novel hum
296	37.5	55.1	504	8	ADS23562	AdS23562 Bacterial	369	37	54.4	353	6	ADB15888	Adb15888 Human PRO
297	37	54.4	20	4	AAE12495	Aae12495 Membrane	370	37	54.4	353	6	ADA47674	Ada47674 Human PRO
298	37	54.4	21	8	ADJ36117	Adj36117 Self-coal	371	37	54.4	353	6	ADA18063	Ada18063 Human sec
299	37	54.4	22	8	ADJ36107	Adj36107 Self-coal	372	37	54.4	353	6	ABO32756	Ab032756 Human sec
300	37	54.4	24	4	AA661232	Aab61232 Human TAN	373	37	54.4	353	6	ADA67469	Ada67469 Human PRO
301	37	54.4	24	6	ABO32669	Ab032669 Secreted	374	37	54.4	353	6	ADB30476	Adb30476 Human PRO
302	37	54.4	24	7	ADB90772	Adb90772 Human TAN	375	37	54.4	353	6	ADA85772	Ada85772 Novel hum
303	37	54.4	24	7	ADF71507	Adf71507 Human TAN	376	37	54.4	353	6	ADA96984	Ada96984 Human PRO
304	37	54.4	24	8	ADQ10325	Adq10325 Human pol	377	37	54.4	353	6	ADA79288	Ada79288 Human PRO
305	37	54.4	73	3	AA240990	Aag240990 Arabidops	378	37	54.4	353	6	ADA87427	Ada87427 Novel hum
306	37	54.4	73	3	AA240990	Aag240990 Arabidops	379	37	54.4	353	6	ADB16629	Adb16629 Human PRO
307	37	54.4	99	4	AAE03217	Aae03217 Human gen	380	37	54.4	353	6	ABO34816	Ab034816 Human PRO
308	37	54.4	99	5	ABG64397	Abg64397 Human alb	381	37	54.4	353	6	ADA16038	Ada16038 Human sec
309	37	54.4	99	6	ADA41244	Ada41244 Human sec	382	37	54.4	353	6	ADA91721	Ada91721 Novel hum
310	37	54.4	99	7	ADC74445	Adc74445 Human sec	383	37	54.4	353	6	ADBL14784	Adb14784 Novel hum
311	37	54.4	99	8	ADL77664	Adl77664 Albumin f	384	37	54.4	353	6	ADBL18745	Adb18745 Novel hum
312	37	54.4	146	4	AA874756	Aab74756 Human sec	385	37	54.4	353	6	ADA93960	Ada93960 Human PRO
313	37	54.4	146	5	ABG65332	Abg65332 Human alb	386	37	54.4	353	6	ADB19856	Adb19856 Novel hum
314	37	54.4	146	8	ADL78599	Adl78599 Albumin f	387	37	54.4	353	6	ADBL13168	Adb13168 Human PRO
315	37	54.4	173	7	ADM26277	Adm26277 Hyperther	388	37	54.4	353	6	ABO43296	Ab043296 Novel hum
316	37	54.4	174	4	AA662210	Aag662210 Human gen	389	37	54.4	353	6	ADA74422	Ada74422 Human PRO
317	37	54.4	174	5	ABG63569	Abg63569 Human alb	390	37	54.4	353	6	ADA42183	Ada42183 Human sec

391	37	54.4	353	6	ADB24655	Human	PRO	Adb24655	Human	PRO	464	37	54.4	353	7	ADC50337	Novel	hum	Adc50337	Novel	hum
392	37	54.4	353	6	ADA82179	Human	PRO	Ada82179	Human	PRO	465	37	54.4	353	7	ADC71884	Novel	hum	Adc71884	Novel	hum
393	37	54.4	353	6	ADA75142	Human	PRO	Ada75142	Human	PRO	466	37	54.4	353	7	ADC59863	Novel	hum	Adc59863	Novel	hum
394	37	54.4	353	6	ADA85220	Novel	hum	Ada85220	Novel	hum	467	37	54.4	353	7	ADC52870	Novel	hum	Adc52870	Novel	hum
395	37	54.4	353	6	ADA84668	Novel	hum	Ada84668	Novel	hum	468	37	54.4	353	7	ADC57224	Novel	hum	Adc57224	Novel	hum
396	37	54.4	353	6	ABO17494	Human	PRO	AbO17494	Human	PRO	469	37	54.4	353	7	ADC60415	Novel	hum	Adc60415	Novel	hum
397	37	54.4	353	6	ADB29924	Human	PRO	Adb29924	Human	PRO	470	37	54.4	353	7	ADC50890	Novel	hum	Adc50890	Novel	hum
398	37	54.4	353	6	ADA80452	Human	PRO	Ada80452	Human	PRO	471	37	54.4	353	7	ADC65417	Human	PRO	Adc65417	Human	PRO
399	37	54.4	353	6	ADA75694	Human	PRO	Ada75694	Human	PRO	472	37	54.4	353	7	ADC54515	Novel	hum	Adc54515	Novel	hum
400	37	54.4	353	6	ADA46919	Human	PRO	Ada46919	Human	PRO	473	37	54.4	353	7	ADC53476	Novel	hum	Adc53476	Novel	hum
401	37	54.4	353	6	ADB25215	Human	PRO	Adb25215	Human	PRO	474	37	54.4	353	7	ADC58999	Novel	hum	Adc58999	Novel	hum
402	37	54.4	353	6	ADA93391	Human	PRO	Ada93391	Human	PRO	475	37	54.4	353	7	ADC55877	Novel	hum	Adc55877	Novel	hum
403	37	54.4	353	6	ADB26741	Human	PRO	Adb26741	Human	PRO	476	37	54.4	353	7	ADC58447	Novel	hum	Adc58447	Novel	hum
404	37	54.4	353	6	ADB31028	Human	PRO	Adb31028	Human	PRO	477	37	54.4	353	7	ADC12246	Human	sec	Adc12246	Human	sec
405	37	54.4	353	6	ADA60956	Homo sapi		Ada60956	Homo sapi		478	37	54.4	353	7	ADD03121	Novel	hum	AdD03121	Novel	hum
406	37	54.4	353	6	ADB24103	Human	PRO	Adb24103	Human	PRO	479	37	54.4	353	7	ADC90113	Novel	hum	Adc90113	Novel	hum
407	37	54.4	353	6	ADA96432	Human	PRO	Ada96432	Human	PRO	480	37	54.4	353	7	ADC69532	Human	PRO	Adc69532	Human	PRO
408	37	54.4	353	6	ADA81004	Human	PRO	Ada81004	Human	PRO	481	37	54.4	353	7	ADC48421	Human	PRO	Adc48421	Human	PRO
409	37	54.4	353	6	ADA95880	Human	PRO	Ada95880	Human	PRO	482	37	54.4	353	7	ADD09950	Human	PRO	AdD09950	Human	PRO
410	37	54.4	353	6	ADB26189	Human	PRO	Adb26189	Human	PRO	483	37	54.4	353	7	ADD04525	Novel	hum	AdD04525	Novel	hum
411	37	54.4	353	6	ADB21674	Novel	hum	Adb21674	Novel	hum	484	37	54.4	353	7	ADC80481	Novel	hum	Adc80481	Novel	hum
412	37	54.4	353	7	ADA77453	Human	PRO	Ada77453	Human	PRO	485	37	54.4	353	7	ADD10988	Human	PRO	AdD10988	Human	PRO
413	37	54.4	353	7	ADB18193	Human	PRO	Adb18193	Human	PRO	486	37	54.4	353	7	ADC47869	Human	PRO	Adc47869	Human	PRO
414	37	54.4	353	7	ADA86876	Novel	hum	Ada86876	Novel	hum	487	37	54.4	353	7	ADD04801	Human	sec	AdD04801	Human	sec
415	37	54.4	353	7	ADA16462	Human	sec	Ada16462	Human	sec	488	37	54.4	353	7	ADC79929	Novel	hum	Adc79929	Novel	hum
416	37	54.4	353	7	ADA12891	Human	sec	Ada12891	Human	sec	489	37	54.4	353	7	ADD09398	Human	PRO	AdD09398	Human	PRO
417	37	54.4	353	7	ADA41759	Human	sec	Ada41759	Human	sec	490	37	54.4	353	7	ADD03807	Human	sec	AdD03807	Human	sec
418	37	54.4	353	7	ADA87979	Novel	hum	Ada87979	Novel	hum	491	37	54.4	353	7	ADD03383	Human	sec	AdD03383	Human	sec
419	37	54.4	353	7	ADA46367	Novel	hum	Ada46367	Novel	hum	492	37	54.4	353	7	ADD41111	Novel	hum	AdD41111	Novel	hum
420	37	54.4	353	7	ADA17106	Human	sec	Ada17106	Human	sec	493	37	54.4	353	7	ADD52250	Human	PRO	AdD52250	Human	PRO
421	37	54.4	353	7	ADA42609	Human	sec	Ada42609	Human	sec	494	37	54.4	353	7	ADD52990	Human	PRO	AdD52990	Human	PRO
422	37	54.4	353	7	ADB28397	Human	PRO	Adb28397	Human	PRO	495	37	54.4	353	7	ADD53542	Novel	hum	AdD53542	Novel	hum
423	37	54.4	353	7	ADB28949	Human	PRO	Adb28949	Human	PRO	496	37	54.4	353	7	ADD51698	Human	PRO	AdD51698	Human	PRO
424	37	54.4	353	7	ADA76901	Human	PRO	Ada76901	Human	PRO	497	37	54.4	353	7	ADD02497	Human	PRO	AdD02497	Human	PRO
425	37	54.4	353	7	ADA88531	Novel	hum	Ada88531	Novel	hum	498	37	54.4	353	7	ADD01931	Human	PRO	AdD01931	Human	PRO
426	37	54.4	353	7	ADA97536	Human	PRO	Ada97536	Human	PRO	499	37	54.4	353	7	ADD54113	Novel	hum	AdD54113	Novel	hum
427	37	54.4	353	7	ADB27293	Human	PRO	Adb27293	Human	PRO	500	37	54.4	353	7	ADD92430	Human	PRO	AdD92430	Human	PRO
428	37	54.4	353	7	ADB22226	Novel	hum	Adb22226	Novel	hum	501	37	54.4	353	7	ADD91326	Human	PRO	AdD91326	Human	PRO
429	37	54.4	353	7	ABO17555	Human	PRO	AbO17555	Human	PRO	502	37	54.4	353	7	ADB03340	Human	PRO	AdB03340	Human	PRO
430	37	54.4	353	7	ADA66917	Human	PRO	Ada66917	Human	PRO	503	37	54.4	353	7	ADB32237	Novel	hum	AdB32237	Novel	hum
431	37	54.4	353	7	ADB22778	Human	PRO	Adb22778	Human	PRO	504	37	54.4	353	7	ADB22169	Human	PRO	AdB22169	Human	PRO
432	37	54.4	353	7	ADB23551	Human	PRO	Adb23551	Human	PRO	505	37	54.4	353	7	ADD79393	Human	PRO	AdD79393	Human	PRO
433	37	54.4	353	7	ADA92273	Novel	hum	Ada92273	Novel	hum	506	37	54.4	353	7	ADB41929	Human	PRO	AdB41929	Human	PRO
434	37	54.4	353	7	ADB15336	Human	PRO	Adb15336	Human	PRO	507	37	54.4	353	7	ADB17746	Human	PRO	AdB17746	Human	PRO
435	37	54.4	353	7	ADB90771	Human	TAN	Adb90771	Human	TAN	508	37	54.4	353	7	ADD91878	Human	PRO	AdD91878	Human	PRO
436	37	54.4	353	7	ADB38588	Novel	hum	Adb38588	Novel	hum	509	37	54.4	353	7	ADB33341	Novel	hum	AdB33341	Novel	hum
437	37	54.4	353	7	ADB38036	Novel	hum	Adb38036	Novel	hum	510	37	54.4	353	7	ADB33893	Novel	hum	AdB33893	Novel	hum
438	37	54.4	353	7	ADB66508	Novel	hum	Adb66508	Novel	hum	511	37	54.4	353	7	ADD79945	Human	PRO	AdD79945	Human	PRO
439	37	54.4	353	7	ADB89588	Human	PRO	Adb89588	Human	PRO	512	37	54.4	353	7	ADD92982	Human	PRO	AdD92982	Human	PRO
440	37	54.4	353	7	ADB90320	Human	PRO	Adb90320	Human	PRO	513	37	54.4	353	7	ADB19402	Human	PRO	AdB19402	Human	PRO
441	37	54.4	353	7	ADB77528	Human	sec	Adb77528	Human	sec	514	37	54.4	353	7	ADB34635	Human	sec	AdB34635	Human	sec
442	37	54.4	353	7	ADB39421	Novel	hum	Adb39421	Novel	hum	515	37	54.4	353	7	ADB18850	Human	PRO	AdB18850	Human	PRO
443	37	54.4	353	7	ADB74664	Human	sec	Adb74664	Human	sec	516	37	54.4	353	7	ADB43046	Human	PRO	AdB43046	Human	PRO
444	37	54.4	353	7	ADB47044	Novel	hum	Adb47044	Novel	hum	517	37	54.4	353	7	ADD95835	Human	PRO	AdD95835	Human	PRO
445	37	54.4	353	7	ADB86651	Human	PRO	Adb86651	Human	PRO	518	37	54.4	353	7	ADB22721	Human	PRO	AdB22721	Human	PRO
446	37	54.4	353	7	ADB77256	Novel	hum	Adb77256	Novel	hum	519	37	54.4	353	7	ADD78839	Human	PRO	AdD78839	Human	PRO
447	37	54.4	353	7	ADB34413	Human	PRO	Adb34413	Human	PRO	520	37	54.4	353	7	ADB32789	Novel	hum	AdB32789	Novel	hum
448	37	54.4	353	7	ADB35517	Human	PRO	Adb35517	Human	PRO	521	37	54.4	353	7	ADB42481	Human	PRO	AdB42481	Human	PRO
449	37	54.4	353	7	ADB33861	Human	PRO	Adb33861	Human	PRO	522	37	54.4	353	7	ADB89525	Human	PRO	AdB89525	Human	PRO
450	37	54.4	353	7	ADB34965	Human	PRO	Adb34965	Human	PRO	523	37	54.4	353	7	ADB40809	Human	PRO	AdB40809	Human	PRO
451	37	54.4	353	7	ADB36069	Human	PRO	Adb36069	Human	PRO	524	37	54.4	353	7	ADB40508	Human	PRO	AdB40508	Human	PRO
452	37	54.4	353	7	ADB46464	Novel	hum	Adb46464	Novel	hum	525	37	54.4	353	7	ADB04508	Human	PRO	AdB04508	Human	PRO
453	37	54.4	353	7	ADC28310	Human	sec	Adb28310	Human	sec	526	37	54.4	353	7	ADB27377	Novel	hum	AdB27377	Novel	hum
454	37	54.4	353	7	ADC39510	Human	sec	Adb39510	Human	sec	527	37	54.4	353	7	ADF71506	Human	TAN	AdF71506	Human	TAN
455	37	54.4	353	7	ADC40024	Human	sec	Adb40024	Human	sec	528	37	54.4	353	7	ADG21446	Novel	hum	AdG21446	Novel	hum
456	37	54.4	353	7	ADC18852	Human	sec	Adb18852	Human	sec	529	37	54.4	353	7	ADG23087	Novel	hum	AdG23087	Novel	hum
457	37	54.4	353	7	ADC34148	Human	sec	Adb34148	Human	sec	530	37	54.4	353	7	ADF97422	Human	PRO	AdF97422	Human	PRO
458	37	54.4	353	7	ADC29203	Human	sec	Adb29203	Human	sec	531	37	54.4	353	7	ADG80486	Human	PRO	AdG80486	Human	PRO
459	37	54.4	353	7	ADC28734	Human	sec	Adb28734	Human	sec	532	37	54.4	353	7	ADG79934	Human	PRO	AdG79934	Human	PRO
460	37	54.4	353	7	ADC40619	Human	sec	Adb40619	Human	sec	533	37	54.4	353	7	ADH59118	Human	sec	AdH59118	Human	sec
461	37	54.4	353</																		

537	37	54.4	353	7	AD1633997	Adi633997 Novel hum	610	37	54.4	353	8	ADG57933	Novel hum
538	37	54.4	353	7	ADI64946	Adi64946 Novel hum	611	37	54.4	353	8	ADG53517	Novel hum
539	37	54.4	353	7	ADI63445	Adi63445 Novel hum	612	37	54.4	353	8	ADG71403	Novel hum
540	37	54.4	353	7	ADH81859	Adh81859 Novel hum	613	37	54.4	353	8	ADG81590	Human PRO
541	37	54.4	353	7	ADH81307	Adh81307 Novel hum	614	37	54.4	353	8	ADH30552	Human PRO
542	37	54.4	353	7	ADJ26165	Adj26165 Novel hum	615	37	54.4	353	8	ADH11919	Novel hum
543	37	54.4	353	7	ADH82476	Adh82476 Novel hum	616	37	54.4	353	8	ADG52341	Novel hum
544	37	54.4	353	7	ADN15875	Adn15875 Novel hum	617	37	54.4	353	8	ADG54069	Novel hum
545	37	54.4	353	7	ADN16504	Adn16504 Novel hum	618	37	54.4	353	8	ADG81038	Human PRO
546	37	54.4	353	7	ADN15323	Adn15323 Novel hum	619	37	54.4	353	8	ADG56277	Novel hum
547	37	54.4	353	7	ADN14771	Adn14771 Novel hum	620	37	54.4	353	8	ADH12543	Novel hum
548	37	54.4	353	8	ADC81033	Adc81033 Novel hum	621	37	54.4	353	8	ADG61389	Novel hum
549	37	54.4	353	8	ADE79080	Ade79080 Human sec	622	37	54.4	353	8	ADH28476	Human PRO
550	37	54.4	353	8	ADD76481	Add76481 Human PRO	623	37	54.4	353	8	ADG54621	Novel hum
551	37	54.4	353	8	ADD87845	Add87845 Human PRO	624	37	54.4	353	8	ADG59661	Novel hum
552	37	54.4	353	8	ADD866249	Add866249 Human PRO	625	37	54.4	353	8	ADH20314	Human sec
553	37	54.4	353	8	ADE79504	Ade79504 Human sec	626	37	54.4	353	8	ADH07169	Human sec
554	37	54.4	353	8	ADE75697	Ade75697 Human PRO	627	37	54.4	353	8	ADH59714	Human sec
555	37	54.4	353	8	ADE73180	Ade73180 Human sec	628	37	54.4	353	8	ADH06742	Human sec
556	37	54.4	353	8	ADE23373	Ade23373 Human PRO	629	37	54.4	353	8	Adi181085	Human PRO
557	37	54.4	353	8	ADE23825	Ade23825 Human PRO	630	37	54.4	353	8	Adi118484	Human sec
558	37	54.4	353	8	ADE24468	Ade24468 Human PRO	631	37	54.4	353	8	Adi165204	Human sec
559	37	54.4	353	8	ADD87293	Add87293 Human PRO	632	37	54.4	353	8	Adi137467	Human sec
560	37	54.4	353	8	ADE89159	Ade89159 Human PRO	633	37	54.4	353	8	ADG09828	Novel hum
561	37	54.4	353	8	ADE73715	Ade73715 Human sec	634	37	54.4	353	8	ADH97271	Human sec
562	37	54.4	353	8	ADE18298	Ade18298 Human PRO	635	37	54.4	353	8	Adi115299	Novel hum
563	37	54.4	353	8	ADE88607	Ade88607 Human PRO	636	37	54.4	353	8	ADG09176	Novel hum
564	37	54.4	353	8	ADE99269	Ade99269 Human sec	637	37	54.4	353	8	Adi165631	Human sec
565	37	54.4	353	8	ADE94627	Ade94627 Human PRO	638	37	54.4	353	8	Adi14631	Novel hum
566	37	54.4	353	8	ADE91038	Ade91038 Human PRO	639	37	54.4	353	8	ADH60374	Human sec
567	37	54.4	353	8	ADE95179	Ade95179 Human PRO	640	37	54.4	353	8	Adi118226	Novel hum
568	37	54.4	353	8	ADE93289	Ade93289 Human PRO	641	37	54.4	353	8	Adi199431	Human sec
569	37	54.4	353	8	ADP34870	Adp34870 Human PRO	642	37	54.4	353	8	ADL08624	Human sec
570	37	54.4	353	8	ADE98388	Ade98388 Human sec	643	37	54.4	353	8	Adm24969	Human sec
571	37	54.4	353	8	ADE92185	Ade92185 Novel hum	644	37	54.4	353	8	Adj163507	Novel hum
572	37	54.4	353	8	ADE90486	Ade90486 Human PRO	645	37	54.4	353	8	Adm29715	Human sec
573	37	54.4	353	8	ADE91633	Ade91633 Novel hum	646	37	54.4	353	8	Adj177402	Human PRO
574	37	54.4	353	8	ADE98815	Ade98815 Human sec	647	37	54.4	353	8	Adj165524	Human PRO
575	37	54.4	353	8	ADG40285	Adg40285 Human sec	648	37	54.4	353	8	Adm27660	Human PRO
576	37	54.4	353	8	ADP73679	Adp73679 Human sec	649	37	54.4	353	8	Adm42384	Human PRO
577	37	54.4	353	8	ADG02212	Adg02212 Human PRO	650	37	54.4	353	8	Adm06037	Human PRO
578	37	54.4	353	8	ADG21998	Adg21998 Novel hum	651	37	54.4	353	8	Adm28246	Human PRO
579	37	54.4	353	8	ADG20068	Adg20068 Human PRO	652	37	54.4	353	8	Adq10324	Human pol
580	37	54.4	353	8	ADP97974	Adp97974 Human PRO	653	37	54.4	353	8	Adr10889	Human sec
581	37	54.4	353	8	ADG24191	Adg24191 Novel hum	654	37	54.4	353	8	Adr17798	Human sec
582	37	54.4	353	8	ADP98545	Adp98545 Human PRO	655	37	54.4	353	8	Adi195728	Human PRO
583	37	54.4	353	8	ADG03376	Adg03376 Human PRO	656	37	54.4	353	8	Adi196280	Novel hum
584	37	54.4	353	8	ADP99097	Adp99097 Human PRO	657	37	54.4	353	8	Adr49418	Human end
585	37	54.4	353	8	ADG16682	Adg16682 Human PRO	658	37	54.4	353	8	Adt03474	Human sec
586	37	54.4	353	8	ADG05141	Adg05141 Human PRO	659	37	54.4	353	8	AdS74437	Human sec
587	37	54.4	353	8	ADG19408	Adg19408 Human PRO	660	37	54.4	360	7	ABO68674	Pseudomon
588	37	54.4	353	8	ADP73255	Adp73255 Human sec	661	37	54.4	366	5	ABb06345	C. biferm
589	37	54.4	353	8	ADG13245	Adg13245 Human PRO	662	37	54.4	371	3	AG311391	Arabidops
590	37	54.4	353	8	ADG08302	Adg08302 Novel hum	663	37	54.4	372	3	AG311390	Arabidops
591	37	54.4	353	8	ADG15472	Adg15472 Human PRO	664	37	54.4	378	7	ABO67443	Klebsiell
592	37	54.4	353	8	ADP96870	Adp96870 Human PRO	665	37	54.4	383	6	ABU47507	Protein e
593	37	54.4	353	8	ADG06055	Adg06055 Human PRO	666	37	54.4	383	6	ABU45474	Protein e
594	37	54.4	353	8	ADG23639	Adg23639 Novel hum	667	37	54.4	384	6	ABU31707	Protein e
595	37	54.4	353	8	ADG03328	Adg03328 Human PRO	668	37	54.4	388	2	AY42435	Cuticle-d
596	37	54.4	353	8	ADG24829	Adg24829 Novel hum	669	37	54.4	403	5	ABP69545	Human pol
597	37	54.4	353	8	ADG07126	Adg07126 Novel hum	670	37	54.4	413	8	ADQ07573	Streptomy
598	37	54.4	353	8	ADG07678	Adg07678 Human PRO	671	37	54.4	420	3	AG311389	Arabidops
599	37	54.4	353	8	ADG55173	Adg55173 Novel hum	672	37	54.4	455	4	AAE03202	Human gen
600	37	54.4	353	8	ADG60837	Adg60837 Novel hum	673	37	54.4	455	5	ABG64398	Human alb
601	37	54.4	353	8	ADG61941	Adg61941 Novel hum	674	37	54.4	455	6	ADA40694	Human sec
602	37	54.4	353	8	ADG92098	Adg92098 Human sec	675	37	54.4	455	7	ADC74085	Human sec
603	37	54.4	353	8	ADG82142	Adg82142 Human PRO	676	37	54.4	455	8	ADL77665	Albumin f
604	37	54.4	353	8	ADG57381	Adg57381 Novel hum	677	37	54.4	592	3	AAy95724	Cosmid cH
605	37	54.4	353	8	ADG56829	Adg56829 Novel hum	678	37	54.4	640	8	ADQ65739	Novel hum
606	37	54.4	353	8	ADG55725	Adg55725 Novel hum	679	37	54.4	654	4	ABB63266	Drosophil
607	37	54.4	353	8	ADG58485	Adg58485 Novel hum	680	37	54.4	661	5	ABG93298	C. albica
608	37	54.4	353	8	ADG70851	Adg70851 Novel hum	681	37	54.4	667	5	ABG63572	Human alb
609	37	54.4	353	8	ADG92525	Adg92525 Human sec	682	37	54.4	683	4	AG62172	Human gen

610	37	54.4	353	8	ADG57933	Novel hum
611	37	54.4	353	8	ADG53517	Novel hum
612	37	54.4	353	8	ADG71403	Novel hum
613	37	54.4	353	8	ADG81590	Human PRO
614	37	54.4	353	8	ADH30552	Human PRO
615	37	54.4	353	8	ADH11919	Novel hum
616	37	54.4	353	8	ADG52341	Novel hum
617	37	54.4	353	8	ADG54069	Novel hum
618	37	54.4	353	8	ADG81038	Human PRO
619	37	54.4	353	8	ADG56277	Novel hum
620	37	54.4	353	8	ADH12543	Novel hum
621	37	54.4	353	8	ADG61389	Novel hum
622	37	54.4	353	8	ADH28476	Human PRO
623	37	54.4	353	8	ADG54621	Novel hum
624	37	54.4	353	8	ADG59661	Novel hum
625	37	54.4	353	8	ADH20314	Human sec
626	37	54.4	353	8	ADH07169	Human sec
627	37	54.4	353	8	ADH59714	Human sec
628	37	54.4	353	8	ADH06742	Human sec
629	37	54.4	353	8	ADI181085	Human PRO
630	37	54.4	353	8	ADI118484	Human sec
631	37	54.4	353	8	ADI165204	Human sec
632	37	54.4	353	8	ADI137467	Human sec
633	37	54.4	353	8	ADG09828	Novel hum
634	37	54.4	353	8	ADH97271	Human sec
635	37	54.4	353	8	ADI115299	Novel hum
636	37	54.4	353	8	ADG09176	Novel hum
637	37	54.4	353	8	ADI165631	Human sec
638	37	54.4	353	8	ADI14631	Novel hum
639	37	54.4	353	8	ADH60374	Human sec
640	37	54.4	353	8	ADI18226	Novel hum
641	37	54.4	353	8	ADI99431	Human sec
642	37	54.4	353	8	ADL08624	Human sec
643	37	54.4	353	8	ADM24969	Human sec
644	37	54.4	353	8	ADJ163507	Novel hum
645	37	54.4	353	8	ADM29715	Human sec
646	37	54.4	353	8	ADJ177402	Human PRO
647	37	54.4	353	8	ADJ165524	Human PRO
648	37	54.4	353	8	ADM27660	Human PRO
649	37	54.4	353	8	ADM42384	Human PRO
650	37	54.4	353	8	ADM06037	Human PRO
651	37	54.4	353	8	ADM28246	Human PRO
652	37	54.4	353	8	ADQL0324	Human pol
653	37	54.4	353	8	ADR10889	Human sec
654	37	54.4	353	8	ADR17798	Human sec
655	37	54.4	353	8	ADI95728	Human PRO
656	37	54.4	353	8	ADI96280	Novel hum
657	37	54.4	353	8	ADR49418	Human end
658	37	54.4	353	8	ADR03474	Human sec
659	37	54.4	353	8	ADG74437	Human sec
660	37	54.4	353	7	ABO68674	Pseudomon
661	37	54.4	366	5	ABR06345	C. biferm
662	37	54.4	371	3	AGG31391	Arabidops
663	37	54.4	372	3	AGG31390	Arabidops
664	37	54.4	378	6	ABO67443	Klebsiell
665	37	54.4	383	6	ABU47507	Protein e
666	37	54.4	384	6	ABU45474	Protein e
667	37	54.4	384	6	ABU31707	Protein e
668	37	54.4	388	2	AY142435	Cuticle-d
669	37	54.4	403	5	ABP69545	Human pol
670	37	54.4	413	8	ADQ07573	Streptomy
671	37	54.4	420	3	AGG31389	Arabidops
672	37	54.4	455	4	AER03202	Human gen
673	37	54.4	455	5	ABG64398	Human alb
674	37	54.4	455	6	ADA0694	Human sec
675	37	54.4	455	7	ADC74085	Human sec
676	37	54.4	455	8	ADL77665	Albumin f
677	37	54.4	592	3	AY95724	Cosmid cH
678	37	54.4	640	8	ADG65739	Novel hum
679	37	54.4	654	4	ABG63266	Drosophil
680	37	54.4	661	5	ABG93298	C. albica
681	37	54.4	667	4	ABG63572	Human alb
682	37	54.4	683	4	AAG62172	Human gen

683	37	54.4	8	ADL76837	Albumin f	RESULT 1
684	37	54.4	4	ABE63269	Drosophil	ABP70214
685	37	54.4	4	ABG00332	Novel hum	ID ABP70234 standard; peptide; 15 AA.
686	37	54.4	699	ADM26342	Hyperther	XX
687	37	54.4	700	AAB95135	Human pro	ABP70234;
688	37	54.4	701	ABU34750	Protein e	XX
689	37	54.4	701	ABU36496	Protein e	DT 07-APR-2003 (first entry)
690	37	54.4	702	AAV14969	Amino aci	XX
691	37	54.4	702	AAB83278	Human FAT	DE Amino acid sequence of targeting factor Gelan094.
692	37	54.4	730	AAV14969	Human PRO	XX
693	37	54.4	730	AAB44255	Human PRO	Lipid-nucleic acid complex; polycation; targeting factor; gene therapy;
694	37	54.4	730	AAB24054	Human PRO	cancer; infection; immune deficiency; gene defect; genetic disease.
695	37	54.4	730	AAB60388	Human fat	XX
696	37	54.4	730	AAB83279	Human FAT	Unidentified.
697	37	54.4	730	ABO25201	Novel hum	XX
698	37	54.4	730	ABU72207	Novel hum	Key
699	37	54.4	730	ADA84115	Human POM	Location/Qualifiers
700	37	54.4	730	ABU84887	Human sec	Modified-site 1
701	37	54.4	730	ABU61085	Human PRO	/note= "s(galactose), cholesteryl-succinyl, DOPE-
702	37	54.4	730	ABU80354	Human sec	succinyl, DSPE-PEG5K-succinyl, DMPE-PEG5K-succinyl, or
703	37	54.4	730	ADA24641	Novel hum	DSPE-PEG5K-succinyl attached"
704	37	54.4	730	ABO19656	Novel hum	WO200288318-A2.
705	37	54.4	730	ADA12302	Human sec	07-NOV-2002.
706	37	54.4	730	ABO19547	Novel hum	30-APR-2002; 2002WO-US013609.
707	37	54.4	730	ADB73608	Human PRO	30-APR-2001; 2001US-0287786P.
708	37	54.4	730	ADB76324	Human PRO	(TARG-) TARGETED GENETICS CORP.
709	37	54.4	730	ADC43750	Human sec	(EMER-) EMERALD GENE SYSTEMS LTD.
710	37	54.4	730	ADC61510	Human sec	✓ Harvie P, Paul R, Cudmore S, O'mahony DJ;
711	37	54.4	730	ADC63474	Human sec	WPI; 2003-183837/18.
712	37	54.4	730	ADC66574	Human sec	Lipid-nucleic acid complex useful for delivering a nucleic acid to a
713	37	54.4	730	ADC68698	Human sec	cell, comprises compacted nucleic acid, polycation, targeting factor and
714	37	54.4	730	ADC62758	Human sec	lipid, and does not comprise protamine or its salt.
715	37	54.4	730	ADC67823	Human sec	Disclosure; Page 42; 259pp; English.
716	37	54.4	730	ADC41143	Human sec	The specification describes a lipid-nucleic acid complex, comprising a
717	37	54.4	730	ADC67198	Human sec	compacted nucleic acid, a polycation, a targeting factor and a lipid, but
718	37	54.4	730	ADC62134	Human sec	not a protamine. The targeting factor increases cellular bioavailability
719	37	54.4	730	ADC41767	Human sec	of the nucleic acid without interaction with a specific outer cell
720	37	54.4	730	ADC49136	Human sec	surface membrane receptor. The mean diameter of the complex is greater
721	37	54.4	730	ADC35190	Human sec	than 100 nm and less than 400 nm. The lipid-nucleic acid complex is
722	37	54.4	730	ADE16304	Human sec	useful for delivering a nucleic acid to a cell in vivo, e.g. for gene
723	37	54.4	730	ADD72919	Human sec	therapy. It reduces levels of inflammatory cytokines such as tumour
724	37	54.4	730	ADD72277	Human sec	CC necrosis factor-alpha. The complex is useful for manufacturing a
725	37	54.4	730	ADE16928	Human sec	medicament for treating or diagnosing a variety of diseases, conditions
726	37	54.4	730	ADF46942	Human sec	or syndromes such as cancer, bacterial, viral or parasitic infections,
727	37	54.4	730	ADG52699	Human sec	CC immune deficiencies, gene defects, and gene deficiencies (e.g. inherited
728	37	54.4	730	ADG60019	Human sec	CC genetic diseases). The present sequence represents a targeting factor for
729	37	54.4	730	ADI60779	Human sec	CC lipid-nucleic acid complexes of the invention
730	37	54.4	730	ADE48436	Human sec	XX
731	37	54.4	730	ADE89537	Human sec	XX
732	37	54.4	730	ADF61177	Human sec	XX
733	37	54.4	730	ADF39869	Human sec	XX
734	37	54.4	730	ADF45665	Human sec	XX
735	37	54.4	730	ADF24061	Human sec	XX
736	37	54.4	730	ADF40493	Human sec	XX
737	37	54.4	730	ADF23437	Human sec	XX
738	37	54.4	730	ADF33420	Human sec	XX
739	37	54.4	730	ADF6887	Human sec	XX
740	37	54.4	730	ADF27523	Human sec	XX
741	37	54.4	730	ADF41117	Human sec	XX
742	37	54.4	730	ADF32796	Human sec	XX
743	37	54.4	730	ADF25162	Human sec	XX
744	37	54.4	730	ADF26263	Human sec	XX
745	37	54.4	730	ADF34052	Human sec	XX
746	37	54.4	730	ADF46289	Human sec	XX
747	37	54.4	730	ADG50275	Human sec	XX
748	37	54.4	730	ADG49651	Human sec	XX
749	37	54.4	730	ADG51523	Human sec	XX
750	37	54.4	730	ADG49027	Human sec	XX

683	37	54.4	8	ADL76837	Albumin f	RESULT 2
684	37	54.4	4	ABE63269	Drosophil	ABP70214
685	37	54.4	4	ABG00332	Novel hum	ID ABP70214 standard; peptide; 15 AA.
686	37	54.4	699	ADM26342	Hyperther	XX
687	37	54.4	700	AAB95135	Human pro	XX
688	37	54.4	701	ABU34750	Protein e	XX
689	37	54.4	701	ABU36496	Protein e	XX
690	37	54.4	702	AAV14969	Amino aci	XX
691	37	54.4	702	AAB83278	Human FAT	XX
692	37	54.4	730	AAV14969	Human PRO	XX
693	37	54.4	730	AAB44255	Human PRO	XX
694	37	54.4	730	AAB24054	Human PRO	XX
695	37	54.4	730	AAB60388	Human fat	XX
696	37	54.4	730	AAB83279	Human FAT	XX
697	37	54.4	730	ABO25201	Novel hum	XX
698	37	54.4	730	ABU72207	Novel hum	XX
699	37	54.4	730	ADA84115	Human POM	XX
700	37	54.4	730	ABU84887	Human sec	XX
701	37	54.4	730	ABU61085	Human PRO	XX
702	37	54.4	730	ABU80354	Human sec	XX
703	37	54.4	730	ADA24641	Novel hum	XX
704	37	54.4	730	ABO19656	Novel hum	XX
705	37	54.4	730	ADA12302	Human sec	XX
706	37	54.4	730	ABO19547	Novel hum	XX
707	37	54.4	730	ADB73608	Human PRO	XX
708	37	54.4	730	ADB76324	Human PRO	XX
709	37	54.4	730	ADC43750	Human sec	XX
710	37	54.4	730	ADC61510	Human sec	XX
711	37	54.4	730	ADC63474	Human sec	XX
712	37	54.4	730	ADC66574	Human sec	XX
713	37	54.4	730	ADC68698	Human sec	XX
714	37	54.4	730	ADC62758	Human sec	XX
715	37	54.4	730	ADC67823	Human sec	XX
716	37	54.4	730	ADC41143	Human sec	XX
717	37	54.4	730	ADC67198	Human sec	XX
718	37	54.4	730	ADC62134	Human sec	XX
719	37	54.4	730	ADC41767	Human sec	XX
720	37	54.4	730	ADC49136	Human sec	XX
721	37	54.4	730	ADC35190	Human sec	XX
722	37	54.4	730	ADE16304	Human sec	XX
723	37	54.4	730	ADD72919	Human sec	XX
724	37	54.4	730	ADD72277	Human sec	XX
725	37	54.4	730	ADE16928	Human sec	XX
726	37	54.4	730	ADF46942	Human sec	XX
727	37	54.4	730	ADG52699	Human sec	XX
728	37	54.4	730	ADG60019	Human sec	XX
729	37	54.4	730	ADI60779	Human sec	XX
730	37	54.4	730	ADE48436	Human sec	XX
731	37	54.4	730	ADE89537	Human sec	XX
732	37	54.4	730	ADF61177	Human sec	XX
733	37	54.4	730	ADF39869	Human sec	XX
734	37	54.4	730	ADF45665	Human sec	XX
735	37	54.4	730	ADF24061	Human sec	XX
736	37	54.4	730	ADF40493	Human sec	XX
737	37	54.4	730	ADF23437	Human sec	XX
738	37	54.4	730	ADF33420	Human sec	XX
739	37	54.4	730	ADF6887	Human sec	XX
740	37	54.4	730	ADF27523	Human sec	XX
741	37	54.4	730	ADF41117	Human sec	XX
742	37	54.4	730	ADF32796	Human sec	XX
743	37	54.4	730	ADF25162	Human sec	XX
744	37	54.4	730	ADF26263	Human sec	XX
745	37	54.4	730	ADF34052	Human sec	XX
746	37	54.4	730	ADF46289	Human sec	XX
747	37	54.4	730	ADG50275	Human sec	XX
748	37	54.4	730	ADG49651	Human sec	XX
749	37	54.4	730	ADG51523	Human sec	XX
750	37	54.4	730	ADG49027	Human sec	XX

ALIGNMENTS

Query Match 100.0%; Score 68; DB 6; Length 15;  
Best Local Similarity 100.0%; Pred. No. 6.2e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 KKAADVLLPVLAAAP 15  
|||||  
Db 1 KKAADVLLPVLAAAP 15

RESULT 2  
ABP70214  
ID ABP70214 standard; peptide; 15 AA.  
XX  
XX ABP70214;

```
XX DT 07-APR-2003 (first entry)
XX DE Amino acid sequence of membrane translocating peptide Elan094.
XX KW Lipid-nucleic acid complex; polycation; targeting factor; gene therapy;
XX KW cancer; infection; immune deficiency; gene defect; genetic disease;
XX KW membrane translocating peptide.
XX OS Unidentified.
XX PN WO200288318-A2.
XX PD 07-NOV-2002.
XX PF 30-APR-2002; 2002WO-US013609.
XX PR 30-APR-2001; 2001US-0287786P.
XX PA (TARG-) TARGETED GENETICS CORP.
XX PA (EMER-) EMERALD GENE SYSTEMS LTD.
XX PI Harvie P, Paul R, Cudmore S, O'mahony DJ;
XX WPI; 2003-183837/18.
XX DR
XX PT Lipid-nucleic acid complex useful for delivering a nucleic acid to a
XX PT cell, comprises compacted nucleic acid, polycation, targeting factor and
XX PT lipid, and does not comprise protamine or its salt.
XX PS Disclosure; Page 10; 259pp; English.
XX CC The specification describes a lipid-nucleic acid complex, comprising a
XX CC compacted nucleic acid, a polycation, a targeting factor and a lipid, but
XX CC not a proline. The targeting factor increases cellular bioavailability
XX CC of the nucleic acid without interaction with a specific outer cell
XX CC surface membrane receptor. The mean diameter of the complex is greater
XX CC than 100 nm and less than 400 nm. The lipid-nucleic acid complex is
XX CC useful for delivering a nucleic acid to a cell in vivo, e.g. for gene
XX CC therapy. It reduces levels of inflammatory cytokines such as tumour
XX CC necrosis factor-alpha. The complex is useful for manufacturing a
XX CC medicament for treating or diagnosing a variety of diseases, conditions
XX CC or syndromes such as cancer, bacterial, viral or parasitic infections,
XX CC immune deficiencies, gene defects, and gene deficiencies (e.g. inherited
XX CC genetic diseases). The present sequence represents a membrane
XX CC translocating peptide, which is used as the targeting factor in lipid-
XX CC nucleic acid complexes of the invention
XX SQ Sequence 15 AA;
Query Match 100.0%; Score 68; DB 6; Length 15;
Best Local Similarity 100.0%; Pred. No. 6.2e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 1 KKAARVLLPVLLAAP 15
Db 1 KKAARVLLPVLLAAP 15
RESULT 3
ABG75426
ID ABG75426 standard; peptide; 15 AA.
XX AC ABG75426;
XX DT 15-APR-2004 (first entry)
XX DE Membrane translocating peptide #1.
XX KW Membrane translocating peptide; transmembrane transport.
XX KW enzymatic degradation; gastrointestinal tract.
XX OS Unidentified.
XX DT 07-APR-2003 (first entry)
XX DE Amino acid sequence of membrane translocating peptide Elan094.
XX KW Lipid-nucleic acid complex; polycation; targeting factor; gene therapy;
XX KW cancer; infection; immune deficiency; gene defect; genetic disease;
XX KW membrane translocating peptide.
XX OS Unidentified.
XX PN WO200288318-A2.
XX PD 07-NOV-2002.
XX PF 30-APR-2002; 2002WO-US013609.
XX PR 30-APR-2001; 2001US-0287786P.
XX PA (TARG-) TARGETED GENETICS CORP.
XX PA (EMER-) EMERALD GENE SYSTEMS LTD.
XX PI Harvie P, Paul R, Cudmore S, O'mahony DJ;
XX WPI; 2003-183837/18.
XX DR
XX PT Lipid-nucleic acid complex useful for delivering a nucleic acid to a
XX PT cell, comprises compacted nucleic acid, polycation, targeting factor and
XX PT lipid, and does not comprise protamine or its salt.
XX PS Disclosure; Page 10; 259pp; English.
XX CC The specification describes a lipid-nucleic acid complex, comprising a
XX CC compacted nucleic acid, a polycation, a targeting factor and a lipid, but
XX CC not a proline. The targeting factor increases cellular bioavailability
XX CC of the nucleic acid without interaction with a specific outer cell
XX CC surface membrane receptor. The mean diameter of the complex is greater
XX CC than 100 nm and less than 400 nm. The lipid-nucleic acid complex is
XX CC useful for delivering a nucleic acid to a cell in vivo, e.g. for gene
XX CC therapy. It reduces levels of inflammatory cytokines such as tumour
XX CC necrosis factor-alpha. The complex is useful for manufacturing a
XX CC medicament for treating or diagnosing a variety of diseases, conditions
XX CC or syndromes such as cancer, bacterial, viral or parasitic infections,
XX CC immune deficiencies, gene defects, and gene deficiencies (e.g. inherited
XX CC genetic diseases). The present sequence represents a membrane
XX CC translocating peptide, which is used as the targeting factor in lipid-
XX CC nucleic acid complexes of the invention
XX SQ Sequence 15 AA;
Query Match 100.0%; Score 68; DB 6; Length 15;
Best Local Similarity 100.0%; Pred. No. 6.2e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 1 KKAARVLLPVLLAAP 15
Db 1 KKAARVLLPVLLAAP 15
RESULT 4
AAU00633
ID AAU00633 standard; peptide; 16 AA.
XX AC AAU00633;
XX DT 07-SEP-2001 (first entry)
XX DE Human membrane translocating peptide (MTLP) #2.
XX KW Membrane translocating peptide; MTLP; human; intracellular gene delivery;
XX KW epithelial cell layer; gastrointestinal tract; circulatory system.
XX OS Homo sapiens.
XX PN WO200127154-A2.
XX PD 19-APR-2001.
XX PF 27-SEP-2000; 2000WO-IB001491.
XX PR 27-SEP-1999; 99US-0156246P.
XX PA (OMAH/) O'MAHONY D J.
```

PA (LAMB/) LAMBKIN I J.  
 XX O'mahony DJ, Lambkin IJ;  
 PI  
 DR WPI; 2001-300212/31.  
 DR N-PSDB; AAS00627.  
 XX  
 CC Compositions for enhancing uptake of e.g. drugs or DNA across a cell  
 PT membrane, comprise membrane translocating peptides having specific amino  
 PT acid sequences or a derivative, fragment, motif, analog or peptidomimetic  
 PT of the peptides.  
 XX  
 XX Claim 2; Page 11; 42pp; English.  
 PS  
 CC The sequence represents a human membrane translocated peptide (MTLP).  
 CC MTLPs and their related fragments, motifs, derivatives and analogues are  
 CC used for enhancing uptake of a pharmaceutically active agent into a cell,  
 CC into or out of an intracellular compartment and across a cell layer (for  
 CC example, an epithelial cell layer lining the gastrointestinal tract),  
 CC either directly or from a pharmaceutically active agent loaded particle,  
 CC into the circulatory system of an animal. This method is useful for  
 CC intracellular gene delivery, as a rapid screening method for the  
 CC identification of MTLPs which retain the functional activity of a full-  
 CC length MTLP, as a cell-based screen for assaying the functional activity  
 CC of a MTLP and characterising the properties of a MTLP, for diagnosis of a  
 CC pathological disorder (by administration of a MTLP-active agent complex  
 CC or MTLP-active particle complex comprising a diagnostic agent) and for  
 CC preventing or treating a pathological disorder  
 XX  
 XX Sequence 16 AA;

Query Match 100.0%; Score 68; DB 4; Length 16;  
 Best Local Similarity 100.0%; Pred. No. 6.6e-05;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKAAAVLLPVLAAAP 15  
 |||||  
 DB 2 KKAAAVLLPVLAAAP 16

RESULT 5  
 ABP70215  
 ID ABP70215 standard; peptide; 16 AA.  
 XX  
 AC ABP70215;  
 XX  
 DT 07-APR-2003 (first entry)  
 XX  
 DE Amino acid sequence of membrane translocating peptide Zelan094.  
 XX  
 KW Lipid-nucleic acid complex; polycation; targeting factor; gene therapy;  
 KW cancer; infection; immune deficiency; gene defect; genetic disease;  
 KW membrane translocating peptide.  
 XX  
 OS Unidentified.  
 XX  
 PN WO200288318-A2.  
 XX  
 PD 07-NOV-2002.  
 XX  
 PF 30-APR-2002; 2002WO-US013609.  
 XX  
 PR 30-APR-2001; 2001US-0287786P.  
 XX  
 XX (TARG-) TARGETED GENETICS CORP.  
 PA (EMER-) EMERALD GENE SYSTEMS LTD.  
 XX  
 PI Harvie P, Paul R, Cudmore S, O'mahony DJ;  
 XX  
 XX WPI; 2003-183837/18.  
 DR  
 XX Lipid-nucleic acid complex useful for delivering a nucleic acid to a  
 PT cell, comprises compacted nucleic acid, polycation, targeting factor and

PT lipid, and does not comprise protamine or its salt.  
 XX  
 PS Disclosure; Page 10; 259pp; English.  
 XX  
 CC The specification describes a lipid-nucleic acid complex, comprising a  
 CC compacted nucleic acid, a polycation, a targeting factor and a lipid, but  
 CC not a protamine. The targeting factor increases cellular bioavailability  
 CC of the nucleic acid without interaction with a specific outer cell  
 CC surface membrane receptor. The mean diameter of the complex is greater  
 CC than 100 nm and less than 400 nm. The lipid-nucleic acid complex is  
 CC useful for delivering a nucleic acid to a cell in vivo, e.g. for gene  
 CC therapy. It reduces levels of inflammatory cytokines such as tumour  
 CC necrosis factor-alpha. The complex is useful for manufacturing a  
 CC medicament for treating or diagnosing a variety of diseases, conditions  
 CC or syndromes such as cancer, bacterial, viral or parasitic infections,  
 CC immune deficiencies, gene defects, and gene deficiencies (e.g. inherited  
 CC genetic diseases). The present sequence represents a membrane  
 CC translocating peptide, which is used as the targeting factor in lipid-  
 CC nucleic acid complexes of the invention  
 XX  
 SQ Sequence 16 AA;

Query Match 100.0%; Score 68; DB 6; Length 16;  
 Best Local Similarity 100.0%; Pred. No. 6.6e-05;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKAAAVLLPVLAAAP 15  
 |||||  
 DB 2 KKAAAVLLPVLAAAP 16

RESULT 6  
 ADB88776  
 ID ADB88776 standard; peptide; 16 AA.  
 XX  
 AC ADB88776;  
 XX  
 DT 04-DEC-2003 (first entry)  
 XX  
 DE Membrane translocating peptide #2.  
 XX  
 KW Peyer's patch cell; non-Peyer's patch cell; transcription factor;  
 KW upregulated protein; antigen; vaccine delivery; M cell;  
 KW membrane translocating peptide.  
 XX  
 OS Unidentified.  
 XX  
 PN WO2003004646-A2.  
 XX  
 PD 16-JAN-2003.  
 XX  
 PF 04-APR-2002; 2002WO-IB003866.  
 XX  
 PR 04-APR-2001; 2001US-0281387P.  
 PR 02-JUL-2001; 2001US-0302591P.  
 XX  
 PA (OMAH/) O'MAHONY D J.  
 XX  
 PI O'mahony DJ, Byrne D, Brayden D;  
 XX  
 XX WPI; 2003-229409/22.  
 DR

Increasing the levels of a protein in a Peyer's patch cell, useful for  
 PT targeted vaccine or drug delivery, comprises delivering to the Peyer's  
 PT patch cell a transcription factor or an activator of a transcription  
 PT factor.

Example 6; Page 51; 147pp; English.

The invention relates to a novel method for increasing the levels of a  
 CC protein in a Peyer's patch cell. The method comprises delivering to the  
 CC cell a nucleic acid coding for a protein, the level of which or its mRNA  
 CC is greater than in a non-Peyer's patch cell. The preferred protein of the

CC invention is a transcription factor or a protein that activates a  
 CC transcription factor selected from Jun-B, c-jun related TF, Jun-D, STAT 3  
 CC -signal transducer and activator of transcription 3, Nfkapb2bgr; Tf p105  
 CC subunit, S-myc proto-oncogene, myc related, Nm23-M2, nucleoside  
 CC diphosphate kinase B, metastasis reducing protein, and C-est-1 proto-  
 CC oncogene, and p54. The preferred upregulated protein of the invention is  
 CC selected from clusterin, T-cell surface glycoprotein CD5 precursor, Hsp  
 CC 84, and Ca2+ dependent phospholipase A2 precursor and the mRNA is for a  
 CC protein selected from the group. The method is useful for increasing or  
 CC decreasing the level of a protein in a Peyer's patch cell, particularly  
 CC in increasing antigen or vaccine delivery to M cells. The method may also  
 CC be used to enhance transport of a drug through the gastrointestinal tract  
 CC (GIT). This sequence represents a membrane translocating peptide of the  
 CC invention.

XX SQ Sequence 16 AA;

Query Match 100.0%; Score 68; DB 7; Length 16;  
 Best Local Similarity 100.0%; Pred. No. 6.6e-05;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKAARAVLLPVLAAAP 15  
 |||||  
 Db 2 KKAARAVLLPVLAAAP 16  
 |||||

#### RESULT 7

AAU00655  
 ID AAU00655 standard; peptide; 19 AA.

XX AC AAU00655;

XX DT 07-SEP-2001 (first entry)

XX DE Human targeting peptide sequence #1.

XX KW Membrane translocating peptide; MTLP; human; intracellular gene delivery;  
 XX KW epithelial cell layer; gastrointestinal tract; circulatory system;  
 XX KW targeting peptide.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

XX FT Modified-site 1

XX FT /note= "N-terminus has a dansyl group"

XX FT Modified-site 19

XX FT /note= "C-terminal amide"

XX PN WO200127154-A2.

XX PD 19-APR-2001.

XX PF 27-SEP-2000; 2000WO-IB001491.

XX PR 27-SEP-1999; 99US-0156246P.

XX PA (OMAH//) O'MAHONY D J.

XX PA (LAMB//) LAMBKIN I J.

XX PI O'mahony DJ, Lambkin IJ;

XX DR WPI; 2001-300212/31.

XX PT Compositions for enhancing uptake of e.g. drugs or DNA across a cell  
 XX PT membrane, comprise membrane translocating peptides having specific amino  
 XX PT acid sequences or a derivative, fragment, motif, analog or peptidomimetic  
 XX PT of the peptides.

XX PS Example 1; Page 22; 42pp; English.

XX CC The sequence represents a human targeting peptide which is used to  
 CC identify the presence of a human membrane translocated peptide (MTLP) and  
 CC quantify the amount, to bind the MTLP to the surface of a particle, or to

CC localise the MTLP in a cell or tissue sample. MTLPs and their related  
 CC fragments, motifs, derivatives and analogues are used for enhancing  
 CC uptake of a pharmaceutically active agent into a cell, into or out of an  
 CC intracellular compartment and across a cell layer (for example, an  
 CC epithelial cell layer lining the gastrointestinal tract), either directly  
 CC or from a pharmaceutically active agent loaded particle, into the  
 CC circulatory system of an animal. This method is useful for intracellular  
 CC gene delivery, as a rapid screening method for the identification of  
 CC MTLPs which retain the functional activity of a full-length MTLP, as a  
 CC cell-based screen for assaying the functional activity of a MTLP and  
 CC characterising the properties of a MTLP, for diagnosis of a pathological  
 CC disorder (by administration of a MTLP-active agent complex or MTLP-active  
 CC treating a pathological disorder

XX SQ Sequence 19 AA;

Query Match 100.0%; Score 68; DB 4; Length 19;  
 Best Local Similarity 100.0%; Pred. No. 8e-05;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKAARAVLLPVLAAAP 15  
 |||||  
 Db 1 KKAARAVLLPVLAAAP 15  
 |||||

#### RESULT 8

AAU00634  
 ID AAU00634 standard; peptide; 19 AA.

XX AC AAU00634;

XX DT 07-SEP-2001 (first entry)

XX DE Human membrane translocating peptide (MTLP) #3.

XX KW Membrane translocating peptide; MTLP; human; intracellular gene delivery;  
 XX KW epithelial cell layer; gastrointestinal tract; circulatory system.

XX OS Homo sapiens.

XX PN WO200127154-A2.

XX PD 19-APR-2001.

XX PF 27-SEP-2000; 2000WO-IB001491.

XX PR 27-SEP-1999; 99US-0156246P.

XX PA (OMAH//) O'MAHONY D J. ✓  
 XX PA (LAMB//) LAMBKIN I J.

XX PI O'mahony DJ, Lambkin IJ;

XX DR WPI; 2001-300212/31.

XX DR N-PSDB; AAS00628.

XX PT Compositions for enhancing uptake of e.g. drugs or DNA across a cell  
 XX PT membrane, comprise membrane translocating peptides having specific amino  
 XX PT acid sequences or a derivative, fragment, motif, analog or peptidomimetic  
 XX PT of the peptides.

XX PS Claim 2; Page 11; 42pp; English.

XX CC The sequence represents a human membrane translocated peptide (MTLP).  
 XX CC MTLPs and their related fragments, motifs, derivatives and analogues are  
 XX CC used for enhancing uptake of a pharmaceutically active agent into a cell,  
 XX CC into or out of an intracellular compartment and across a cell layer (for  
 XX CC example, an epithelial cell layer lining the gastrointestinal tract),  
 XX CC either directly or from a pharmaceutically active agent loaded particle,  
 XX CC into the circulatory system of an animal. This method is useful for  
 XX CC intracellular gene delivery, as a rapid screening method for the  
 XX CC identification of MTLPs which retain the functional activity of a full-

CC length MTLp, as a cell-based screen for assaying the functional activity  
 CC of a MTLp and characterising the properties of a MTLp, for diagnosis of a  
 CC pathological disorder (by administration of a MTLp-active agent complex  
 CC or MTLp-active particle complex comprising a diagnostic agent) and for  
 CC preventing or treating a pathological disorder  
 XX  
 SQ Sequence 19 AA;

Query Match 100.0%; Score 68; DB 4; Length 19;  
 Best Local Similarity 100.0%; Pred. No. 8e-05;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKAARVLLPVLAAAP 15  
 DB 1 KKAARVLLPVLAAAP 15

RESULT 9  
 ADB88777  
 ID ADB88777 standard; peptide; 19 AA.

AC ADB88777;  
 DT 04-DEC-2003 (first entry)  
 XX Membrane translocating peptide #3.  
 DE  
 XX Peyer's patch cell; non-Peyer's patch cell; transcription factor;  
 KW upregulated protein; antigen; vaccine delivery; M cell;  
 KW membrane translocating peptide.  
 XX  
 OS Unidentified.  
 XX

PN WO2003004646-A2.

PD 16-JAN-2003.

PF 04-APR-2002; 2002WO-IB003866.

PR 04-APR-2001; 2001US-0281387P.

PR 02-JUL-2001; 2001US-0302591P.

PA (OMAH/ O'MAHONY D J.

PI O'mahony DJ, Byrne D, Brayden D;

XX WPI; 2003-229409/22.

XX Increasing the levels of a protein in a Peyer's patch cell, useful for  
 PT targeted vaccine or drug delivery, comprises delivering to the Peyer's  
 PT patch cell a transcription factor or an activator of a transcription  
 PT factor.

PS Example 6; Page 51; 147pp; English.

XX The invention relates to a novel method for increasing the levels of a  
 CC protein in a Peyer's patch cell. The method comprises delivering to the  
 CC cell a nucleic acid coding for a protein, the level of which or its mRNA  
 CC is greater than in a non-Peyer's patch cell. The preferred protein of the  
 CC invention is a transcription factor or a protein that activates a  
 CC transcription factor selected from Jun-B, c-Jun related TF, Jun-D, STAT 3  
 CC -signal transducer and activator of transcription 3, NFkappaBgr; Tf p105  
 CC subunit, S-myc proto-oncogene, myc related, Nm23-M2, nucleoside  
 CC diphosphate kinase B, metacaspase reducing protein, and C-est-1 proto-  
 CC oncogene, and p54. The preferred upregulated protein of the invention is  
 CC selected from clusterin, T-cell surface glycoprotein CD5 precursor, HSP  
 CC 84, and Ca2+ dependent phospholipase A2 precursor and the mRNA is for a  
 CC protein selected from the group. The method is useful for increasing or  
 CC decreasing the level of a protein in a Peyer's patch cell, particularly  
 CC in increasing antigen or vaccine delivery to M cells. The method may also  
 CC be used to enhance transport of a drug through the gastrointestinal tract  
 CC (GIT). This sequence represents a membrane translocating peptide of the  
 CC invention.

XX  
 SQ Sequence 19 AA;  
 Query Match 100.0%; Score 68; DB 7; Length 19;  
 Best Local Similarity 100.0%; Pred. No. 8e-05;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKAARVLLPVLAAAP 15  
 DB 1 KKAARVLLPVLAAAP 15

RESULT 10  
 ABG75428  
 ID ABG75428 standard; peptide; 19 AA.

AC ABG75428;

DT 15-APR-2004 (first entry)

DE Membrane translocating peptide #3.

KW Membrane translocating peptide; transmembrane transport;  
 KW enzymatic degradation; gastrointestinal tract.

OS Unidentified.

FH Key Location/Qualifiers  
 FT modified\_site 2 /note= "modified by epsilon-dansyl"  
 FT modified\_site 19 /note= "C-terminal amide"

PN WO2003089458-A2.

PD 30-OCT-2003.

PF 21-APR-2003; 2003WO-US012543.

PR 19-APR-2002; 2002US-00126845.

PA (SARL-) SARLAN LTD.

PI O'mahony DJ, Lambkin I, Houghten R, Pinilla C;

XX WPI; 2004-034528/03.

XX Composition, useful for treating a pathological disease in an animal,  
 PT comprises a translocating peptide, consisting of a transport peptide, an  
 PT extended peptide comprising the transport peptide or a transport-active  
 PT fragment.

PS Claim 1; Page 24; Opp; English.

XX The present invention relates to a composition which comprises a  
 CC translocating peptide, consisting of a transport peptide, an extended  
 CC peptide comprising the transport peptide or a transport-active fragment  
 CC of at least 4 amino acids of the transport peptide. The composition is  
 CC useful in treating a pathological disorder in an animal, as it protects  
 CC an orally delivered active agent from enzymatic degradation in the  
 CC gastrointestinal tract (GIT), and promotes absorption across epithelial  
 CC cells lining the GIT. The present sequence is a polypeptide used in the  
 CC exemplification of the invention

XX Sequence 19 AA;

Query Match 100.0%; Score 68; DB 8; Length 19;  
 Best Local Similarity 100.0%; Pred. No. 8e-05;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKAARVLLPVLAAAP 15  
 DB 1 KKAARVLLPVLAAAP 15



RESULT 11  
ABP70216  
ID ABP70216 standard; peptide; 20 AA.  
XX  
AC ABP70216;  
XX  
DT 07-APR-2003 (first entry)  
XX  
DE Amino acid sequence of membrane translocating peptide ZELAN094R.  
XX  
KW Lipid-nucleic acid complex; polycation; targeting factor; gene therapy;  
KW cancer; infection; immune deficiency; gene defect; genetic disease;  
KW membrane translocating peptide.  
XX  
OS Unidentified.  
XX  
PN WO200288318-A2.  
XX  
PD 07-NOV-2002.  
XX  
PF 30-APR-2002; 2002WO-US013609.  
XX  
PR 30-APR-2001; 2001US-0287786P.  
XX  
PA (TARG-) TARGETED GENETICS CORP.  
PA (EMER-) EMERALD GENE SYSTEMS LTD.  
XX  
PI Harvie P, Paul R, Cudmore S, O'Mahony-Bor,  
XX  
DR WPI; 2003-183837/18.  
XX  
PT Lipid-nucleic acid complex useful for delivering a nucleic acid to a  
PT cell, comprises compacted nucleic acid, polycation, targeting factor and  
PT lipid, and does not comprise protamine or its salt.  
XX  
PS Disclosure; Page 10; 259pp; English.  
XX  
PF The specification describes a lipid-nucleic acid complex, comprising a  
CC compacted nucleic acid, a polycation, a targeting factor and a lipid, but  
CC not a protamine. The targeting factor increases cellular bioavailability  
CC of the nucleic acid without interaction with a specific outer cell  
CC surface membrane receptor. The mean diameter of the complex is greater  
CC than 100 nm and less than 400 nm. The lipid-nucleic acid complex is  
CC useful for delivering a nucleic acid to a cell in vivo, e.g. for gene  
CC therapy. It reduces levels of inflammatory cytokines such as tumour  
CC necrosis factor-alpha. The complex is useful for manufacturing a  
CC medicament for treating or diagnosing a variety of diseases, conditions  
CC or syndromes such as cancer, bacterial, viral or parasitic infections,  
CC immune deficiencies, gene defects, and gene deficiencies (e.g. inherited  
CC genetic diseases). The present sequence represents a membrane  
CC translocating peptide, which is used as the targeting factor in lipid-  
CC nucleic acid complexes of the invention  
XX  
SQ Sequence 20 AA;  
Query Match 100.0%; Score 68; DB 6; Length 20;  
Best Local Similarity 100.0%; Pred. No. 8.5e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 KKAQAVLLPVLLAAP 15  
Db 2 KKAQAVLLPVLLAAP 16  
RESULT 12  
AAU00632  
ID AAU00632 standard; peptide; 21 AA.  
XX  
AC AAU00632;  
XX  
DT 07-SEP-2001 (first entry)

XX Human membrane translocating peptide (MTLP) #1.  
DE Membrane translocating peptide; MTLP; human; intracellular gene delivery;  
XX epithelial cell layer; gastrointestinal tract; circulatory system.  
KW Homo sapiens.  
XX  
PN WO200127154-A2.  
XX  
PD 19-APR-2001.  
XX  
PF 27-SEP-2000; 2000WO-IB001491.  
XX  
PR 27-SEP-1999; 99US-0156246P.  
XX  
PA (OMAH/) O'MAHONY D J.  
PA (LAMB/) LAMBKIN I J.  
XX  
PI O'mahony DJ, Lambkin IJ;  
XX  
DR WPI; 2001-300212/31.  
DR N-PSDB; AAS00626.  
XX  
PT Compositions for enhancing uptake of e.g. drugs or DNA across a cell  
PT membrane, comprise membrane translocating peptides having specific amino  
PT acid sequences or a derivative, fragment, motif, analog or peptidomimetic  
PT of the peptides.  
XX  
PS Claim 1; Page 11; 42pp; English.  
XX  
CC The sequence represents a human membrane translocated peptide (MTLP).  
CC MTLPs and their related fragments, motifs, derivatives and analogues are  
CC used for enhancing uptake of a pharmaceutically active agent into a cell,  
CC into or out of an intracellular compartment and across a cell layer (for  
CC example, an epithelial cell layer lining the gastrointestinal tract),  
CC either directly or from a pharmaceutically active agent loaded particle,  
CC into the circulatory system of an animal. This method is useful for  
CC intracellular gene delivery, as a rapid screening method for the  
CC identification of MTLPs which retain the functional activity of a full-  
CC length MTLP, as a cell-based screen for assaying the functional activity  
CC of a MTLP and characterising the properties of a MTLP, for diagnosis of a  
CC pathological disorder (by administration of a MTLP-active agent complex  
CC or MTLP-active particle comprising a diagnostic agent) and for  
CC preventing or treating a pathological disorder  
XX  
SQ Sequence 21 AA;  
Query Match 100.0%; Score 68; DB 4; Length 21;  
Best Local Similarity 100.0%; Pred. No. 9e-05;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 KKAQAVLLPVLLAAP 15  
Db 1 KKAQAVLLPVLLAAP 15  
RESULT 13  
ADB88775  
ID ADB88775 standard; peptide; 21 AA.  
XX  
AC ADB88775;  
XX  
DT 04-DEC-2003 (first entry)  
XX  
DE Membrane translocating peptide #1.  
KW Peyer's patch cell; non-Peyer's patch cell; transcription factor;  
KW upregulated protein; antigen; vaccine delivery; M cell;  
XX membrane translocating peptide.  
XX  
OS Unidentified.

PN WO2003004646-A2.  
 XX 16-JAN-2003.  
 XX 04-APR-2002; 2002WO-IB003866.  
 XX 04-APR-2001; 2001US-0281387P.  
 PR 02-JUL-2001; 2001US-0302591P.  
 XX (OMAH/) O'MAHONY D J.  
 PA O'mahony DJ, Byrne D, Brayden D;  
 PI WPI; 2003-229409/22.  
 XX Increasing the levels of a protein in a Peyer's patch cell, useful for  
 PT targeted vaccine or drug delivery, comprises delivering to the Peyer's  
 PT patch cell a transcription factor or an activator of a transcription  
 PT factor.  
 XX Example 6; Page 51; 147pp; English.  
 XX The invention relates to a novel method for increasing the levels of a  
 CC protein in a Peyer's patch cell. The method comprises delivering to the  
 CC cell a nucleic acid coding for a protein, the level of which or its mRNA  
 CC is greater than in a non-Peyer's patch cell. The preferred protein of the  
 CC invention is a transcription factor or a protein that activates a  
 CC transcription factor selected from Jun-B, c-Jun related TF, Jun-D, STAT 3  
 CC -signal transducer and activator of transcription 3, NFkappaBgr; Tf p105  
 CC subunit, S-myc proto-oncogene, myc related, Nm23-M2, nucleoside  
 CC diphosphate kinase B, metastasis reducing protein, and C-sest-I proto-  
 CC oncogene, and p54. The preferred upregulated protein of the invention is  
 CC selected from clusterin, T-cell surface glycoprotein CD5 precursor, HSP  
 CC 84, and Ca2+-dependent phospholipase A2 precursor and the mRNA is for a  
 CC protein selected from the group. The method is useful for increasing or  
 CC decreasing the level of a protein in a Peyer's patch cell, particularly  
 CC in increasing antigen or vaccine delivery to M cells. The method may also  
 CC be used to enhance transport of a drug through the gastrointestinal tract  
 CC (GIT). This sequence represents a membrane translocating peptide of the  
 CC invention.  
 XX Sequence 21 AA;  
 SQ  
 Query Match 100.0%; Score 68; DB 7; Length 21;  
 Best Local Similarity 100.0%; Pred. No. 9e-05;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 KKAAAVLLPVLAAAP 15  
 |||||  
 DB 1 KKAAAVLLPVLAAAP 15  
 |||||  
 RESULT 14  
 AAU00635  
 ID AAU00635 standard; peptide; 16 AA.  
 XX AAU00635;  
 AC  
 XX 07-SEP-2001 (first entry)  
 DT  
 XX Human membrane translocating peptide (MTLP) #4.  
 DE  
 XX Membrane translocating peptide; MTLP; human; intracellular gene delivery;  
 KW epithelial cell layer; gastrointestinal tract; circulatory system.  
 XX Homo sapiens.  
 OS  
 XX WO200127154-A2.  
 FN  
 XX 19-APR-2001.  
 PD  
 XX 27-SEP-2000; 2000WO-IB001491.  
 PF  
 XX

PR 27-SEP-1999; 99US-0156246P.  
 XX (OMAH/) O'MAHONY D J. ✓  
 PA (LAMB/) LAMBKIN I J.  
 XX O'mahony DJ, Lambkin IJ;  
 FI  
 XX WPI; 2001-300212/31.  
 DR N-PSDB; AAS00629.  
 DR  
 XX Compositions for enhancing uptake of e.g. drugs or DNA across a cell  
 PT membrane, comprise membrane translocating peptides having specific amino  
 PT acid sequences or a derivative, fragment, motif, analog or peptidomimetic  
 PT of the peptides.  
 XX Claim 2; Page 11; 42pp; English.  
 PS  
 XX The sequence represents a human membrane translocated peptide (MTLP).  
 CC MTLPs and their related fragments, motifs, derivatives and analogues are  
 CC used for enhancing uptake of a pharmaceutically active agent into a cell,  
 CC into or out of an intracellular compartment and across a cell layer (for  
 CC example, an epithelial cell layer lining the gastrointestinal tract),  
 CC either directly or from a pharmaceutically active agent loaded particle,  
 CC into the circulatory system of an animal. This method is useful for  
 CC identification of MTLPs which retain the functional activity of a full-  
 CC length MTLP, as a cell-based screen for assaying the functional activity of a full-  
 CC intracellular gene delivery, as a rapid screening method for the  
 CC of a MTLP and characterising the properties of a MTLP, for diagnosis of a  
 CC pathological disorder (by administration of a MTLP-active agent complex  
 CC or MTLP-active particle complex comprising a diagnostic agent) and for  
 CC preventing or treating a pathological disorder  
 XX Sequence 16 AA;  
 SQ  
 Query Match 94.1%; Score 64; DB 4; Length 16;  
 Best Local Similarity 93.3%; Pred. No. 0.00033;  
 Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 KKAAAVLLPVLAAAP 15  
 |||||  
 DB 1 KKCAAVLLPVLAAAP 15  
 |||||  
 RESULT 15  
 ADB88778  
 ID ADB88778 standard; peptide; 16 AA.  
 XX ADB88778;  
 AC  
 XX 04-DEC-2003 (first entry)  
 DT  
 XX Membrane translocating peptide #4.  
 DE  
 XX Peyer's patch cell; non-Peyer's patch cell; transcription factor;  
 KW upregulated protein; antigen; vaccine delivery; M cell;  
 KW membrane translocating peptide.  
 XX Unidentified.  
 OS  
 XX WO2003004646-A2.  
 FN  
 XX 16-JAN-2003.  
 PD  
 XX 04-APR-2002; 2002WO-IB003866.  
 PF  
 XX 04-APR-2001; 2001US-0281387P.  
 PR 02-JUL-2001; 2001US-0302591P.  
 PR  
 XX (OMAH/) O'MAHONY D J.  
 PA O'mahony DJ, Byrne D, Brayden D;  
 PI WPI; 2003-229409/22.  
 PF  
 XX

XX Increasing the levels of a protein in a Peyer's patch cell, useful for  
PT targeted vaccine or drug delivery, comprises delivering to the Peyer's  
PT patch cell a transcription factor or an activator of a transcription  
PT factor.  
XX  
PS Example 6; Page 51; 147pp; English.  
XX  
CC The invention relates to a novel method for increasing the levels of a  
CC protein in a Peyer's patch cell. The method comprises delivering to the  
CC cell a nucleic acid coding for a protein, the level of which or its mRNA  
CC is greater than in a non-Peyer's patch cell. The preferred protein of the  
CC invention is a transcription factor or a protein that activates a  
CC transcription factor selected from Jun-B, c-jun related TF, Jun-D, STAT 3  
CC -signal transducer and activator of transcription 3, Nfkapabgr; Tf p105  
CC subunit, s-myc proto-oncogene, myc related, Nm23-M2, nucleoside  
CC diphosphate kinase B, metastasis reducing protein, and C-est-I proto-  
CC onogene, and p54. The preferred upregulated protein of the invention is  
CC selected from clusterin, T-cell surface glycoprotein CD5 precursor, Hsp  
CC 84, and Ca2+ dependent phospholipase A2 precursor and the mRNA is for a  
CC protein selected from the group. The method is useful for increasing or  
CC decreasing the level of a protein in a Peyer's patch cell, particularly  
CC in increasing antigen or vaccine delivery to M cells. The method may also  
CC be used to enhance transport of a drug through the gastrointestinal tract  
CC (GIT). This sequence represents a membrane translocating peptide of the  
CC invention.  
XX  
SQ Sequence 16 AA;

Query Match 94.1%; Score 64; DB 7; Length 16;  
Best Local Similarity 93.3%; Pred. No. 0.00033;  
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 KKCAAVLLPVLLAAP 15  
||| ||||| |||||  
DB 1 KKCAAVLLPVLLAAP 15

RESULT 16  
ABG75429  
ID ABG75429 standard; peptide; 16 AA.  
XX  
AC ABG75429;  
XX  
DT 15-APR-2004 (first entry)  
XX  
DE Membrane translocating peptide #4.  
XX  
KW Membrane translocating peptide; transmembrane transport;  
KW enzymatic degradation; gastrointestinal tract.  
XX  
OS Unidentified.  
XX  
FH Key Location/Qualifiers  
FT modified\_site 2  
FT /note= "modified by epsilon-dansyl"  
FT modified\_site 16  
FT /note= "C-terminal amide"  
XX  
PN WO2003089458-A2.  
XX  
PD 30-OCT-2003.  
XX  
PF 21-APR-2003; 2003WO-US012543.  
XX  
PR 19-APR-2002; 2002US-00126845.  
XX  
PA (SARL-) SARLAN LTD.  
XX  
PI O'mahony DJ, Lambkin I, Houghten R, Pinilla C;  
XX  
DR WPI; 2004-034528/03.

PT Composition, useful for treating a pathological disease in an animal,  
PT comprises a translocating peptide, consisting of a transport peptide, an  
PT extended peptide comprising the transport peptide or a transport-active  
PT fragment.  
XX  
PS Claim 1; Page 24; Opp; English.  
XX  
CC The present invention relates to a composition which comprises a  
CC translocating peptide, consisting of a transport peptide, an extended  
CC peptide comprising the transport peptide or a transport-active fragment  
CC of at least 4 amino acids of the transport peptide. The composition is  
CC useful in treating a pathological disorder in an animal, as it protects  
CC an orally delivered active agent from enzymatic degradation in the  
CC gastrointestinal tract (GIT), and promotes absorption across epithelial  
CC cells lining the GIT. The present sequence is a polypeptide used in the  
CC exemplification of the invention  
XX  
SQ Sequence 16 AA;  
Query Match 94.1%; Score 64; DB 8; Length 16;  
Best Local Similarity 93.3%; Pred. No. 0.00033;  
Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 KKAAAVLLPVLLAAP 15  
||| ||||| |||||  
DB 1 KKCAAVLLPVLLAAP 15

RESULT 17  
ABG75427  
ID ABG75427 standard; peptide; 14 AA.  
XX  
AC ABG75427;  
XX  
DT 15-APR-2004 (first entry)  
XX  
DE Membrane translocating peptide #2.  
XX  
KW Membrane translocating peptide; transmembrane transport;  
KW enzymatic degradation; gastrointestinal tract.  
XX  
OS Unidentified.  
XX  
FH Key Location/Qualifiers  
FT modified\_site 14  
FT /note= "C-terminal amide"  
XX  
PN WO2003089458-A2.  
XX  
PD 30-OCT-2003.  
XX  
PF 21-APR-2003; 2003WO-US012543.  
XX  
PR 19-APR-2002; 2002US-00126845.  
XX  
PA (SARL-) SARLAN LTD.  
XX  
PI O'mahony DJ, Lambkin I, Houghten R, Pinilla C;  
XX  
DR WPI; 2004-034528/03.  
XX  
PT Composition, useful for treating a pathological disease in an animal,  
PT comprises a translocating peptide, consisting of a transport peptide, an  
PT extended peptide comprising the transport peptide or a transport-active  
PT fragment.  
XX  
PS Claim 1; Page 24; Opp; English.  
XX  
CC The present invention relates to a composition which comprises a  
CC translocating peptide, consisting of a transport peptide, an extended  
CC peptide comprising the transport peptide or a transport-active fragment  
CC of at least 4 amino acids of the transport peptide. The composition is  
CC useful in treating a pathological disorder in an animal, as it protects  
CC an orally delivered active agent from enzymatic degradation in the  
CC gastrointestinal tract (GIT), and promotes absorption across epithelial  
CC cells lining the GIT. The present sequence is a polypeptide used in the  
CC exemplification of the invention  
XX  
SQ Sequence 16 AA;

CC an orally delivered active agent from enzymatic degradation in the  
 CC gastrointestinal tract (GIT), and promotes absorption across epithelial  
 CC cells lining the GIT. The present sequence is a polypeptide used in the  
 CC exemplification of the invention  
 XX  
 SQ Sequence 14 AA;

Query Match 92.6%; Score 63; DB 8; Length 14;  
 Best Local Similarity 100.0%; Pred. No. 0.00042;  
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 KAAAVLLPVLAAAP 15  
 DB 1 KAAAVLLPVLAAAP 14

RESULT 18  
 ABP70217  
 ID ABP70217 standard; peptide; 15 AA.

AC ABP70217;  
 XX  
 DT 07-APR-2003 (first entry)  
 XX  
 DE Amino acid sequence of membrane translocating peptide ZELAN207.

XX Lipid-nucleic acid complex; polycation; targeting factor; gene therapy;  
 KW cancer; infection; immune deficiency; gene defect; genetic disease;  
 KW membrane translocating peptide.

XX Unidentified.  
 XX  
 PN WO200288318-A2.

XX 07-NOV-2002.  
 XX  
 PD 30-APR-2002; 2002WO-US013609.  
 XX  
 PF 30-APR-2001; 2001US-0287786P.

XX (TARG-) TARGETED GENETICS CORP.  
 PA (EMER-) EMERALD GENE SYSTEMS LTD.

XX Harvie P, Paul R, Cudmore S, O'mahony DJ;

XX WPI; 2003-183837/18.

XX Lipid-nucleic acid complex useful for delivering a nucleic acid to a  
 PT cell, comprises compacted nucleic acid, polycation, targeting factor and  
 PT lipid, and does not comprise protamine or its salt.

XX Disclosure; Page 10; 259pp; English.

XX The specification describes a lipid-nucleic acid complex, comprising a  
 CC compacted nucleic acid, a polycation, a targeting factor and a lipid, but  
 CC not a protamine. The targeting factor increases cellular bioavailability  
 CC of the nucleic acid without interaction with a specific outer cell  
 CC surface membrane receptor. The mean diameter of the complex is greater  
 CC than 100 nm and less than 400 nm. The lipid-nucleic acid complex is  
 CC useful for delivering a nucleic acid to a cell in vivo, e.g. for gene  
 CC therapy. It reduces levels of inflammatory cytokines such as tumour  
 CC necrosis factor-alpha. The complex is useful for manufacturing a  
 CC medicament for treating or diagnosing a variety of diseases, conditions  
 CC or syndromes such as cancer, bacterial, viral or parasitic infections,  
 CC immune deficiencies, gene defects, and gene deficiencies (e.g. inherited  
 CC genetic diseases). The present sequence represents a membrane  
 CC translocating peptide, which is used as the targeting factor in lipid-  
 CC nucleic acid complexes of the invention

XX Sequence 15 AA;

Query Match 92.6%; Score 63; DB 6; Length 15;  
 Best Local Similarity 93.3%; Pred. No. 0.00046;

Matches 14; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 KKAARVLLPVLAAAP 15  
 DB 1 KKAARVLLPVLAAAP 15

RESULT 19  
 ABP70233  
 ID ABP70233 standard; peptide; 15 AA.

XX ABP70233;  
 AC  
 DT 07-APR-2003 (first entry)  
 XX

XX Amino acid sequence of targeting factor Elan094G.

XX Lipid-nucleic acid complex; polycation; targeting factor; gene therapy;  
 KW cancer; infection; immune deficiency; gene defect; genetic disease.

XX Unidentified.

XX Key Location/Qualifiers  
 FH Modified-site 1  
 FT /note= "Des-Pro attached"  
 FT Modified-site 15  
 FT /note= "galactose attached"

XX WO200288318-A2.

XX 07-NOV-2002.

XX 30-APR-2002; 2002WO-US013609.

XX 30-APR-2001; 2001US-0287786P.

XX (TARG-) TARGETED GENETICS CORP.  
 PA (EMER-) EMERALD GENE SYSTEMS LTD.

XX Harvie P, Paul R, Cudmore S, O'mahony DJ;

XX WPI; 2003-183837/18.

XX Lipid-nucleic acid complex useful for delivering a nucleic acid to a  
 PT cell, comprises compacted nucleic acid, polycation, targeting factor and  
 PT lipid, and does not comprise protamine or its salt.

XX Disclosure; Page 42; 259pp; English.

XX The specification describes a lipid-nucleic acid complex, comprising a  
 CC compacted nucleic acid, a polycation, a targeting factor and a lipid, but  
 CC not a protamine. The targeting factor increases cellular bioavailability  
 CC of the nucleic acid without interaction with a specific outer cell  
 CC surface membrane receptor. The mean diameter of the complex is greater  
 CC than 100 nm and less than 400 nm. The lipid-nucleic acid complex is  
 CC useful for delivering a nucleic acid to a cell in vivo, e.g. for gene  
 CC therapy. It reduces levels of inflammatory cytokines such as tumour  
 CC necrosis factor-alpha. The complex is useful for manufacturing a  
 CC medicament for treating or diagnosing a variety of diseases, conditions  
 CC or syndromes such as cancer, bacterial, viral or parasitic infections,  
 CC immune deficiencies, gene defects, and gene deficiencies (e.g. inherited  
 CC genetic diseases). The present sequence represents a targeting factor for  
 CC lipid-nucleic acid complexes of the invention

XX Sequence 15 AA;

Query Match 89.7%; Score 61; DB 6; Length 15;  
 Best Local Similarity 100.0%; Pred. No. 0.001;

Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKAARVLLPVLAA 14  
 DB 1 KKAARVLLPVLAA 14

## RESULT 20

ADK50911  
ID ADK50911 standard; peptide; 21 AA.

XX AC ADK50911;

XX DT 04-NOV-2004 (first entry)

XX DE Human C35 / Kaposi FGF signal sequence MST cell-penetrating peptide 123.

XX KW C35 epitope; cytostatic; vaccine; tumour; breast; bladder carcinoma;

XX KW human; Kaposi FGF signal sequence h region MST; cell-penetrating;

XX KW fibroblast growth factor; membrane-translocating signal.

XX OS Homo sapiens.

XX OS Unidentified.

XX PN WO2003104428-A2.

XX PD 18-DEC-2003.

XX PF 10-JUN-2003; 2003WO-US018252.

XX PR 10-JUN-2002; 2002US-0386738P.

XX PR 11-DEC-2002; 2002US-0432241P.

XX PR 23-APR-2003; 2003US-0464650P.

XX PA (VACC-) VACCINEX INC.

XX PA (UYRP ) UNIV ROCHESTER.

XX PI Zauderer M, Evans EE, Borrello MA;

XX DR WPI; 2004-062349/06.

XX PT Novel C35 polypeptide useful for formulation of immunogenic composition

XX PT to induce antibodies and cell-mediated immunity against tumor cells.

XX PS Disclosure; Page 223; 626pp; English.

XX CC The invention relates to a novel isolated polypeptide comprising or  
consisting of two or more C35 peptide epitopes. The polypeptide of the  
invention demonstrates cytostatic activity and may be useful for the  
formulation of an immunogenic composition, such as a vaccine, to induce  
antibodies and cell-mediated immunity against target cells such as tumour  
cells. Furthermore, the polypeptide and its analogues may be useful as  
prognostic markers for carcinoma, such as human breast or bladder  
carcinoma. The current sequence is that of human C35 / Kaposi FGF  
(fibroblast growth factor) signal sequence MST (membrane-translocating  
sequence) cell-penetrating polypeptide of the invention.

XX CC Query Match 85.3%; Score 58; DB 8; Length 21;

XX CC Best Local Similarity 100.0%; Pred.No. 0.0049;

XX CC Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX QY 3 AAALLPVLAAAP 15

XX DB 9 AAALLPVLAAAP 21

XX RESULT 21

XX ADK50919

XX ID ADK50919 standard; peptide; 21 AA.

XX AC ADK50919;

XX DT 04-NOV-2004 (first entry)

XX DE Human C35 / Kaposi FGF signal sequence MST cell-penetrating peptide 131.

XX KW C35 epitope; cytostatic; vaccine; tumour; breast; bladder carcinoma;

XX KW human; Kaposi FGF signal sequence h region MST; cell-penetrating;

XX KW fibroblast growth factor; membrane-translocating signal.

XX OS Homo sapiens.

XX OS Unidentified.

XX PN WO2003104428-A2.

XX PD 18-DEC-2003.

XX PF 10-JUN-2003; 2003WO-US018252.

KW C35 epitope; cytostatic; vaccine; tumour; breast; bladder carcinoma;  
KW human; Kaposi FGF signal sequence h region MST; cell-penetrating;  
XX fibroblast growth factor; membrane-translocating signal.

XX OS Homo sapiens.

XX OS Unidentified.

XX PN WO2003104428-A2.

XX PD 18-DEC-2003.

XX PF 10-JUN-2003; 2003WO-US018252.

XX PR 10-JUN-2002; 2002US-0386738P.

XX PR 11-DEC-2002; 2002US-0432241P.

XX PR 23-APR-2003; 2003US-0464650P.

XX PA (VACC-) VACCINEX INC.

XX PA (UYRP ) UNIV ROCHESTER.

XX PI Zauderer M, Evans EE, Borrello MA;

XX DR WPI; 2004-062349/06.

XX PT Novel C35 polypeptide useful for formulation of immunogenic composition

XX PT to induce antibodies and cell-mediated immunity against tumor cells.

XX PS Disclosure; Page 225; 626pp; English.

XX CC The invention relates to a novel isolated polypeptide comprising or  
consisting of two or more C35 peptide epitopes. The polypeptide of the  
invention demonstrates cytostatic activity and may be useful for the  
formulation of an immunogenic composition, such as a vaccine, to induce  
antibodies and cell-mediated immunity against target cells such as tumour  
cells. Furthermore, the polypeptide and its analogues may be useful as  
prognostic markers for carcinoma, such as human breast or bladder  
carcinoma. The current sequence is that of human C35 / Kaposi FGF  
(fibroblast growth factor) signal sequence MST (membrane-translocating  
sequence) cell-penetrating polypeptide of the invention.

XX CC Query Match 85.3%; Score 58; DB 8; Length 21;

XX CC Best Local Similarity 100.0%; Pred.No. 0.0049;

XX CC Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX QY 3 AAALLPVLAAAP 15

XX DB 9 AAALLPVLAAAP 21

XX RESULT 22

XX ADK50915

XX ID ADK50915 standard; peptide; 21 AA.

XX AC ADK50915;

XX DT 04-NOV-2004 (first entry)

XX DE Human C35 / Kaposi FGF signal sequence MST cell-penetrating peptide 127.

XX KW C35 epitope; cytostatic; vaccine; tumour; breast; bladder carcinoma;

XX KW human; Kaposi FGF signal sequence h region MST; cell-penetrating;

XX KW fibroblast growth factor; membrane-translocating signal.

XX OS Homo sapiens.

XX OS Unidentified.

XX PN WO2003104428-A2.

XX PD 18-DEC-2003.

XX PF 10-JUN-2003; 2003WO-US018252.

XX 10-JUN-2002; 2002US-0386738P.  
 PR 11-DEC-2002; 2002US-0432241P.  
 PR 23-APR-2003; 2003US-0464650P.  
 XX (VACC-) VACCINEX INC.  
 PA (UYRP ) UNIV ROCHESTER.  
 XX Zauderer M, Evans EE, Borrello MA;  
 XX WPI; 2004-062349/06.  
 DR Novel C35 polypeptide useful for formulation of immunogenic composition  
 XX to induce antibodies and cell-mediated immunity against tumor cells.  
 XX Disclosure; Page 224; 626pp; English.  
 XX The invention relates to a novel isolated polypeptide comprising or  
 CC consisting of two or more C35 peptide epitopes. The polypeptide of the  
 CC invention demonstrates cytostatic activity and may be useful for the  
 CC formulation of an immunogenic composition, such as a vaccine, to induce  
 CC antibodies and cell-mediated immunity against target cells such as tumour  
 CC cells. Furthermore, the polypeptide and its analogues may be useful as  
 CC prognostic markers for carcinoma, such as human breast or bladder  
 CC carcinoma. The current sequence is that of human C35 / Kaposi FGF  
 CC (fibroblast growth factor) signal sequence MST (membrane-translocating  
 CC sequence) cell-penetrating polypeptide of the invention.  
 XX Sequence 21 AA;  
 SQ Query Match 85.3%; Score 58; DB 8; Length 21;  
 Best Local Similarity 100.0%; Pred. No. 0.0049;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 3 AAALLPVLAAAP 15  
 Db |||||  
 9 AAALLPVLAAAP 21  
 RESULT 23  
 ADK50916  
 ID ADK50916 standard; peptide; 33 AA.  
 AC ADK50916;  
 XX 04-NOV-2004 (first entry)  
 DT Human C35 / Kaposi FGF signal sequence MST cell-penetrating peptide 128.  
 DE C35 epitope; cytostatic; vaccine; tumour; breast; bladder carcinoma;  
 XX human; Kaposi FGF signal sequence h region MST; cell-penetrating;  
 KW fibroblast growth factor; membrane-translocating signal.  
 XX Homo sapiens.  
 OS Unidentified.  
 XX WO2003104428-A2.  
 PN 18-DEC-2003.  
 PD 10-JUN-2003; 2003WO-US018252.  
 XX 10-JUN-2002; 2002US-0386738P.  
 PR 11-DEC-2002; 2002US-0432241P.  
 PR 23-APR-2003; 2003US-0464650P.  
 XX (VACC-) VACCINEX INC.  
 PA (UYRP ) UNIV ROCHESTER.  
 XX Zauderer M, Evans EE, Borrello MA;  
 XX WPI; 2004-062349/06.  
 DR Novel C35 polypeptide useful for formulation of immunogenic composition  
 XX to induce antibodies and cell-mediated immunity against tumor cells.  
 XX Disclosure; Page 224; 626pp; English.  
 XX The invention relates to a novel isolated polypeptide comprising or  
 CC consisting of two or more C35 peptide epitopes. The polypeptide of the  
 CC invention demonstrates cytostatic activity and may be useful for the  
 CC formulation of an immunogenic composition, such as a vaccine, to induce  
 CC antibodies and cell-mediated immunity against target cells such as tumour  
 CC cells. Furthermore, the polypeptide and its analogues may be useful as  
 CC prognostic markers for carcinoma, such as human breast or bladder  
 CC carcinoma. The current sequence is that of human C35 / Kaposi FGF  
 CC (fibroblast growth factor) signal sequence MST (membrane-translocating  
 CC sequence) cell-penetrating polypeptide of the invention.  
 XX Sequence 21 AA;  
 SQ Query Match 85.3%; Score 58; DB 8; Length 21;  
 Best Local Similarity 100.0%; Pred. No. 0.0049;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

PT Novel C35 polypeptide useful for formulation of immunogenic composition  
 XX to induce antibodies and cell-mediated immunity against tumor cells.  
 XX Disclosure; Page 224; 626pp; English.  
 XX The invention relates to a novel isolated polypeptide comprising or  
 CC consisting of two or more C35 peptide epitopes. The polypeptide of the  
 CC invention demonstrates cytostatic activity and may be useful for the  
 CC formulation of an immunogenic composition, such as a vaccine, to induce  
 CC antibodies and cell-mediated immunity against target cells such as tumour  
 CC cells. Furthermore, the polypeptide and its analogues may be useful as  
 CC prognostic markers for carcinoma, such as human breast or bladder  
 CC carcinoma. The current sequence is that of human C35 / Kaposi FGF  
 CC (fibroblast growth factor) signal sequence MST (membrane-translocating  
 CC sequence) cell-penetrating polypeptide of the invention.  
 XX Sequence 33 AA;  
 SQ Query Match 85.3%; Score 58; DB 8; Length 33;  
 Best Local Similarity 100.0%; Pred. No. 0.0081;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 3 AAALLPVLAAAP 15  
 Db |||||  
 21 AAALLPVLAAAP 33  
 RESULT 24  
 ADK50912  
 ID ADK50912 standard; peptide; 33 AA.  
 AC ADK50912;  
 XX 04-NOV-2004 (first entry)  
 DT Human C35 / Kaposi FGF signal sequence MST cell-penetrating peptide 124.  
 DE C35 epitope; cytostatic; vaccine; tumour; breast; bladder carcinoma;  
 KW human; Kaposi FGF signal sequence h region MST; cell-penetrating;  
 KW fibroblast growth factor; membrane-translocating signal.  
 XX Homo sapiens.  
 OS Unidentified.  
 XX WO2003104428-A2.  
 PN 18-DEC-2003.  
 PD 10-JUN-2003; 2003WO-US018252.  
 XX 10-JUN-2002; 2002US-0386738P.  
 PR 11-DEC-2002; 2002US-0432241P.  
 PR 23-APR-2003; 2003US-0464650P.  
 XX (VACC-) VACCINEX INC.  
 PA (UYRP ) UNIV ROCHESTER.  
 XX Zauderer M, Evans EE, Borrello MA;  
 XX WPI; 2004-062349/06.  
 DR Novel C35 polypeptide useful for formulation of immunogenic composition  
 XX to induce antibodies and cell-mediated immunity against tumor cells.  
 XX Disclosure; Page 223; 626pp; English.  
 XX The invention relates to a novel isolated polypeptide comprising or  
 CC consisting of two or more C35 peptide epitopes. The polypeptide of the  
 CC invention demonstrates cytostatic activity and may be useful for the  
 CC formulation of an immunogenic composition, such as a vaccine, to induce  
 CC antibodies and cell-mediated immunity against target cells such as tumour  
 CC cells. Furthermore, the polypeptide and its analogues may be useful as  
 CC prognostic markers for carcinoma, such as human breast or bladder

CC carcinoma. The current sequence is that of human C35 / Kaposi FGF  
 CC (fibroblast growth factor) signal sequence MST (membrane-translocating  
 CC sequence) cell-penetrating polypeptide of the invention.

XX Sequence 33 AA;

Query Match 85.3%; Score 58; DB 8; Length 33;  
 Best Local Similarity 100.0%; Pred. No. 0.0081;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 AAALLPVLAAAP 15  
 |||||  
 Db 21 AAALLPVLAAAP 33

# RESULT 25

ADK50920  
 ID ADK50920 standard; peptide; 33 AA.

XX AC ADK50920;

XX DT 04-NOV-2004 (first entry)

XX DE Human C35 / Kaposi FGF signal sequence MST cell-penetrating peptide 132.

XX KW C35 epitope; cytostatic; vaccine; tumour; breast; bladder carcinoma;  
 XX human; Kaposi FGF signal sequence h region MST; cell-penetrating;  
 KW fibroblast growth factor; membrane-translocating signal.

XX OS Homo sapiens.  
 OS Unidentified.

PN WO2003104428-A2.

XX PD 18-DEC-2003.

XX PF 10-JUN-2003; 2003WO-US018252.

XX PR 10-JUN-2002; 2002US-0386738P.

XX PR 11-DEC-2002; 2002US-0432241P.

XX PR 23-APR-2003; 2003US-0464650P.

XX PA (VACC-) VACCINEX INC.

XX PA (UVRP ) UNIV ROCHESTER.

XX PI Zauderer M, Evans EE, Borrello MA;

XX DR WPI; 2004-062349/06.

XX PT Novel C35 polypeptide useful for formulation of immunogenic composition

XX PT to induce antibodies and cell-mediated immunity against tumor cells.  
 XX Disclosure; Page 225; 626pp; English.

XX The invention relates to a novel isolated polypeptide comprising or  
 CC consisting of two or more C35 peptide epitopes. The polypeptide of the  
 CC invention demonstrates cytostatic activity and may be useful for the  
 CC formulation of an immunogenic composition, such as a vaccine, to induce  
 CC antibodies and cell-mediated immunity against target cells such as tumour  
 CC cells. Furthermore, the polypeptide and its analogues may be useful as  
 CC prognostic markers for carcinoma, such as human breast or bladder  
 CC carcinoma. The current sequence is that of human C35 / Kaposi FGF  
 CC (fibroblast growth factor) signal sequence MST (membrane-translocating  
 CC sequence) cell-penetrating polypeptide of the invention.

XX SQ Sequence 33 AA;

Query Match 85.3%; Score 58; DB 8; Length 33;  
 Best Local Similarity 100.0%; Pred. No. 0.0081;  
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 AAALLPVLAAAP 15  
 |||||

Db 21 AAALLPVLAAAP 33

# RESULT 26

ADR40510  
 ID ADR40510 standard; peptide; 18 AA.

XX AC ADR40510;

XX DT 18-NOV-2004 (first entry)

XX DE Peptide comprising a membrane translocating sequence #2.

XX KW peptidomimetic; Cytostatic; Tumour; Apoptosis; oncological disorder.

XX OS Unidentified.

PN WO2004073650-A2.

XX PD 02-SEP-2004.

XX PF 20-FEB-2004; 2004WO-US005030.

XX PR 20-FEB-2003; 2003US-0319960P.

XX PA (UYSF-) UNIV SOUTH FLORIDA.

XX PI Turkson J, Jove R, Sebti SM, Hamilton AD;

XX DR WPI; 2004-642398/62.

XX PT New peptidomimetic compounds are tumor cell replication inhibitors useful  
 XX to treat tumors or oncological disorders of e.g. breast, kidney, mouth,  
 PT larynx, esophagus, stomach, testis, cervix, head, neck, colon, ovary,  
 PT lung, bladder and skin.

XX PS Example 5; SEQ ID NO 2; 47pp; English.

XX CC The present invention relates to peptidomimetic compounds and their  
 CC salts. The compound is useful in the treatment of a tumour or an  
 CC oncological disorder (breast, kidney, mouth, larynx, esophagus, stomach,  
 CC testis, cervix, head, neck, colon, ovary, bladder, skin, muscle,  
 CC pancreas, prostate, bone, eye, blood cells or brain cancers) in a human  
 CC or animal. The present sequence represents is a peptide comprising a  
 CC membrane translocating sequence.

XX SQ Sequence 18 AA;

Query Match 83.8%; Score 57; DB 8; Length 18;

Best Local Similarity 86.7%; Pred. No. 0.0062;

Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KKAALLPVLAAAP 15

Db 4 KTKAALLPVLAAAP 18

# RESULT 27

ADR40509

ID ADR40509 standard; peptide; 18 AA.

XX AC ADR40509;

XX DT 18-NOV-2004 (first entry)

XX DE Peptide comprising a membrane translocating sequence #1.

XX KW peptidomimetic; Cytostatic; Tumour; Apoptosis; oncological disorder.

XX OS Unidentified.

XX FH Key Location/Qualifiers

XX Modified-site 2

FT XX WO2004073650-A2. /note= "phosphorylated"

PN XX

XX XX

PD XX 02-SEP-2004.

XX XX

PF XX 20-FEB-2004; 2004WO-US005030.

XX XX

PR XX 20-FEB-2003; 2003US-0319960P.

XX XX

PA (UYSF-) UNIV SOUTH FLORIDA.

XX XX

PI Turkeon J, Jove R, Sebt SM, Hamilton AD;

XX XX

DR WPI; 2004-642398/62.

XX XX

PT New peptidomimetic compounds are tumor cell replication inhibitors useful

PT to treat tumors or oncological disorders of e.g. breast, kidney, mouth,

PT larynx, esophagus, stomach, testis, cervix, head, neck, colon, ovary,

PT lung, bladder and skin.

XX XX

PS Example 3; SEQ ID NO 1; 47pp; English.

XX XX

CC The present invention relates to peptidomimetic compounds and their

CC salts. The compound is useful in the treatment of a tumour or an

CC oncological disorder (breast, kidney, mouth, larynx, esophagus, stomach,

CC testis, cervix, head, neck, colon, ovary, lung, bladder, skin, muscle,

CC pancreas, prostate, bone, eye, blood cells or brain cancers) in a human

CC or animal. The present sequence represents a peptide comprising a

CC membrane translocating sequence.

XX XX

SQ Sequence 18 AA;

Query Match 83.8%; Score 57; DB 8; Length 18;

Best Local Similarity 86.7%; Pred. No. 0.0062;

Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 KKAAVLLPVLAAAP 15

Db 4 KTKAAVLLPVLAAAP 18

RESULT 28

ADKS0831

ID ADKS0831 standard; peptide; 21 AA.

XX AC

XX ADKS0831;

XX XX

DT 04-NOV-2004 (first entry)

XX XX

DE Human C35 / Kaposi FGF signal sequence MST cell-penetrating peptide 43.

XX XX

C35 epitope; cytostatic; vaccine; tumour; breast; bladder carcinoma;

KW human; Kaposi FGF signal sequence h region MST; cell-penetrating;

KW fibroblast growth factor; membrane-translocating signal.

XX XX

OS Homo sapiens.

OS Unidentified.

XX XX

PN WO2003104428-A2.

XX XX

PD 18-DEC-2003.

XX XX

PF 10-JUN-2003; 2003WO-US018252.

XX XX

C35 epitope; cytostatic; vaccine; tumour; breast; bladder carcinoma;

KW human; Kaposi FGF signal sequence h region MST; cell-penetrating;

KW fibroblast growth factor; membrane-translocating signal.

XX XX

OS Homo sapiens.

OS Unidentified.

XX XX

PN WO2003104428-A2.

XX XX

PD 18-DEC-2003.

XX XX

PF 10-JUN-2003; 2003WO-US018252.

XX XX

10-JUN-2002; 2002US-0386738P.

PR 11-DEC-2002; 2002US-0432241P.

PR 23-APR-2003; 2003US-0464650P.

XX XX

PA (VACC-) VACCINEX INC.

PA (UYRP ) UNIV ROCHESTER.

XX XX

PI Zauderer M, Evans EE, Borrello MA;

XX WPI; 2004-062349/06.

DR XX

XX XX

PT Novel C35 polypeptide useful for formulation of immunogenic composition

PT to induce antibodies and cell-mediated immunity against tumor cells.

XX XX

PS Disclosure; Page 203; 626pp; English.

XX XX

CC The invention relates to a novel isolated polypeptide comprising or

CC consisting of two or more C35 peptide epitopes. The polypeptide of the

CC invention demonstrates cytostatic activity and may be useful for the

CC formulation of an immunogenic composition, such as a vaccine, to induce

CC antibodies and cell-mediated immunity against target cells such as tumour

CC cells. Furthermore, the polypeptide and its analogues may be useful as

CC prognostic markers for carcinoma, such as human breast or bladder

CC carcinoma. The current sequence is that of human C35 / Kaposi FGF

CC (fibroblast growth factor) signal sequence MST (membrane-translocating

CC sequence) cell-penetrating polypeptide of the invention.

XX XX

SQ Sequence 21 AA;

Query Match 83.8%; Score 57; DB 8; Length 21;

Best Local Similarity 86.7%; Pred. No. 0.0073;

Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 KKAAVLLPVLAAAP 15

Db 7 KDLAAVLLPVLAAAP 21

RESULT 29

ADKS0832

ID ADKS0832 standard; peptide; 33 AA.

XX AC

XX ADKS0832;

XX XX

DT 04-NOV-2004 (first entry)

XX XX

DE Human C35 / Kaposi FGF signal sequence MST cell-penetrating peptide 44.

XX XX

C35 epitope; cytostatic; vaccine; tumour; breast; bladder carcinoma;

KW human; Kaposi FGF signal sequence h region MST; cell-penetrating;

KW fibroblast growth factor; membrane-translocating signal.

XX XX

OS Homo sapiens.

OS Unidentified.

XX XX

PN WO2003104428-A2.

XX XX

PD 18-DEC-2003.

XX XX

PF 10-JUN-2003; 2003WO-US018252.

XX XX

10-JUN-2002; 2002US-0386738P.

PR 11-DEC-2002; 2002US-0432241P.

PR 23-APR-2003; 2003US-0464650P.

XX XX

PA (VACC-) VACCINEX INC.

PA (UYRP ) UNIV ROCHESTER.

XX XX

PI Zauderer M, Evans EE, Borrello MA;

XX XX

DR WPI; 2004-062349/06.

XX XX

PT Novel C35 polypeptide useful for formulation of immunogenic composition

PT to induce antibodies and cell-mediated immunity against tumor cells.

XX XX

PS Disclosure; Page 203; 626pp; English.

XX XX

CC The invention relates to a novel isolated polypeptide comprising or

CC consisting of two or more C35 peptide epitopes. The polypeptide of the

CC invention demonstrates cytostatic activity and may be useful for the

CC formulation of an immunogenic composition, such as a vaccine, to induce



CC antibodies and cell-mediated immunity against target cells such as tumour  
 CC cells. Furthermore, the polypeptide and its analogues may be useful as  
 CC prognostic markers for carcinoma, such as human breast or bladder  
 CC carcinoma. The current sequence is that of human C35 / Kaposi FGF  
 CC (fibroblast growth factor) signal sequence MST (membrane-translocating  
 CC sequence) cell-penetrating polypeptide of the invention.  
 XX  
 XX SQ Sequence 33 AA;

Query Match 83.8%; Score 57; DB 8; Length 33;  
 Best Local Similarity 86.7%; Pred. No. 0.012; 2; Indels 0; Gaps 0;  
 Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 KKAQAVLLPVLLAAP 15  
 : : | | | | | | | | | |  
 Db 19 KDLAVALLPVLLAAP 33

RESULT 30  
 ADK50791  
 ID ADK50791 standard; peptide; 21 AA.  
 AC ADK50791;

DT 04-NOV-2004 (first entry)

Human C35 / Kaposi FGF signal sequence MST cell-penetrating peptide 3.

C35 epitope; cytostatic; vaccine; tumour; breast; bladder carcinoma;  
 human; Kaposi FGF signal sequence h region MST; cell-penetrating;  
 fibroblast growth factor; membrane-translocating signal.

OS Homo sapiens.  
 OS Unidentified.

PN WO2003104428-A2.

XX 18-DEC-2003.

XX 10-JUN-2003; 2003WO-US018252.

XX 10-JUN-2002; 2002US-0386738P.

PR 11-DEC-2002; 2002US-0432241P.

PR 23-APR-2003; 2003US-0464650P.

XX (VACC-) VACCINEX INC.

PA (UYRP) UNIV ROCHESTER.

PI Zauderer M, Evans EE, Borrello MA;

XX WPI; 2004-062349/06.

XX Novel C35 polypeptide useful for formulation of immunogenic composition

PT to induce antibodies and cell-mediated immunity against tumor cells.

XX Disclosure; Page 193; 626pp; English.

CC The invention relates to a novel isolated polypeptide comprising or  
 CC consisting of two or more C35 peptide epitopes. The polypeptide of the  
 CC invention demonstrates cytostatic activity and may be useful for the  
 CC formulation of an immunogenic composition, such as a vaccine, to induce  
 CC antibodies and cell-mediated immunity against target cells such as tumour  
 CC cells. Furthermore, the polypeptide and its analogues may be useful as  
 CC prognostic markers for carcinoma, such as human breast or bladder  
 CC carcinoma. The current sequence is that of human C35 / Kaposi FGF  
 CC (fibroblast growth factor) signal sequence MST (membrane-translocating  
 CC sequence) cell-penetrating polypeptide of the invention.  
 XX

SQ Sequence 21 AA;

Query Match 82.4%; Score 56; DB 8; Length 21;  
 Best Local Similarity 80.0%; Pred. No. 0.011;  
 Matches 12; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KKAQAVLLPVLLAAP 15  
 : : | | | | | | | | | |  
 Db 7 EEVAAVLLPVLLAAP 21

RESULT 31

ADK50792  
 ID ADK50792 standard; peptide; 33 AA.

XX AC ADK50792;

XX DT 04-NOV-2004 (first entry)

XX Human C35 / Kaposi FGF signal sequence MST cell-penetrating peptide 4.

XX C35 epitope; cytostatic; vaccine; tumour; breast; bladder carcinoma;  
 human; Kaposi FGF signal sequence h region MST; cell-penetrating;  
 fibroblast growth factor; membrane-translocating signal.

OS Homo sapiens.  
 OS Unidentified.

PN WO2003104428-A2.

XX 18-DEC-2003.

XX 10-JUN-2003; 2003WO-US018252.

XX 10-JUN-2002; 2002US-0386738P.

PR 11-DEC-2002; 2002US-0432241P.

PR 23-APR-2003; 2003US-0464650P.

XX (VACC-) VACCINEX INC.

PA (UYRP) UNIV ROCHESTER.

PI Zauderer M, Evans EE, Borrello MA;

XX WPI; 2004-062349/06.

XX Novel C35 polypeptide useful for formulation of immunogenic composition

PT to induce antibodies and cell-mediated immunity against tumor cells.

XX Disclosure; Page 193; 626pp; English.

XX The invention relates to a novel isolated polypeptide comprising or  
 CC consisting of two or more C35 peptide epitopes. The polypeptide of the  
 CC invention demonstrates cytostatic activity and may be useful for the  
 CC formulation of an immunogenic composition, such as a vaccine, to induce  
 CC antibodies and cell-mediated immunity against target cells such as tumour  
 CC cells. Furthermore, the polypeptide and its analogues may be useful as  
 CC prognostic markers for carcinoma, such as human breast or bladder  
 CC carcinoma. The current sequence is that of human C35 / Kaposi FGF  
 CC (fibroblast growth factor) signal sequence MST (membrane-translocating  
 CC sequence) cell-penetrating polypeptide of the invention.  
 XX

SQ Sequence 33 AA;

Query Match 82.4%; Score 56; DB 8; Length 33;  
 Best Local Similarity 80.0%; Pred. No. 0.018;  
 Matches 12; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 KKAQAVLLPVLLAAP 15  
 : : | | | | | | | | | |  
 Db 19 EEVAAVLLPVLLAAP 33

RESULT 32

ADS17601  
 ID ADS17601 standard; peptide; 17 AA.

XX AC ADS17601;

DT 02-DEC-2004 (first entry)  
 XX Membrane transporter peptide from Kaposi fibroblast growth factor.  
 DE  
 XX immunostimulatory; membrane transport; homophilic; signaling protein;  
 KW caspase; kinase; phosphatase; viral protein; tumour antigen;  
 KW nuclear protein; nucleolar protein; DNA synthesis; cytoskeletal protein;  
 KW cell proliferation; cytotaxis; membrane transporter peptide;  
 KW Kaposi fibroblast factor; TAT peptide; HIV-1; antenapedia homeodomain;  
 KW herpes virus protein VP22; transportan peptide; Alzheimer's disease;  
 KW Huntington's disease; Parkinson's disease; antibody.  
 XX  
 OS Synthetic.  
 XX  
 XX WO2004078146-A2.  
 PN  
 XX  
 XX 16-SEP-2004.  
 PD  
 XX  
 XX 05-MAR-2004; 2004WO-US006911.  
 PF  
 XX  
 XX 05-MAR-2003; 2003US-0451980P.  
 PR  
 XX  
 XX (INNE-) INNEKUS BIOTECHNOLOGY INC.  
 PA (IMMP-) IMMPPERON INC.  
 XX  
 XX Kohler H, Muller S, Brown TL, Zhao Y, Morgan AC;  
 PI  
 XX  
 XX WPI; 2004-653567/63.  
 DR  
 XX  
 XX New compound for regulating normal or infected cell function comprising  
 PT an antibody conjugated to a membrane transporter peptide, useful in  
 PT preparing a composition for treating or preventing human diseases, e.g.  
 PT Alzheimer's disease.  
 PT  
 XX  
 PS Claim 8; SEQ ID NO 8; 50pp; English.  
 XX  
 XX The specification describes a fusion protein for regulating normal or  
 CC infected cell function, comprising an antibody conjugated to a peptide  
 CC having immunostimulatory, membrane transport, and homophilic activities.  
 CC The antibody is immunospecific for a signaling protein internal cell  
 CC consisting of caspases, kinases or phosphatases, an immature viral  
 CC protein, a cell-surface or intracellular tumour antigen, a nuclear or  
 CC nucleolar protein participating in regulation of DNA synthesis and gene  
 CC expression, or a cytoskeletal protein participating in cell proliferation  
 CC or cytotaxis. The peptide portion of the fusion protein is preferably a  
 CC membrane transporter peptide that is endogenous to Kaposi fibroblast  
 CC factor, TAT peptides of HIV-1, antenapedia homeodomain-derived peptide,  
 CC herpes virus protein VP22, or transportan peptide. Fusion protein of the  
 CC invention are useful for preparing a composition for treating or  
 CC preventing human diseases, e.g., Alzheimer's disease, Huntington's  
 CC disease or Parkinson's disease. The present sequence represents a  
 CC membrane transporter peptide which is derived from Kaposi fibroblast  
 CC growth factor. The peptide is used to produce fusion proteins of the  
 CC invention.  
 CC  
 XX Sequence 17 AA;  
 SQ  
 Query Match 80.9%; Score 55; DB 8; Length 17;  
 Best Local Similarity 85.7%; Pred. No. 0.013;  
 Matches 12; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 2 KAAAVLLPVLLAAP 15  
 Db : |||||  
 3 EGAAVLLPVLLAAP 16  
 RESULT 33  
 ADR88404  
 ID ADR88404 standard; peptide; 17 AA.  
 XX  
 AC ADR88404;  
 XX  
 DT 16-DEC-2004 (first entry)

XX Mouse MTS membrane translocation peptide.  
 DE  
 XX Mouse; antibody; autophilic; MTS; T15; dimerisation; apoptosis;  
 KW tryptophan photoactivation; malignancy; cancer; tumour;  
 KW auto-immune disorder; transplantation rejection; Alzheimer's disease;  
 KW neurodegenerative condition; membrane translocation peptide.  
 XX  
 OS Mus sp.  
 XX  
 XX US2004185039-A1.  
 PN  
 XX 23-SEP-2004.  
 PD  
 XX  
 XX 29-AUG-2003; 2003US-00652864.  
 PF  
 XX  
 XX 30-AUG-2002; 2002US-0407421P.  
 PR  
 XX  
 XX (KOHL/) KOHLER H.  
 PA (MORG/) MORGAN A C.  
 XX  
 XX Kohler H, Morgan AC;  
 PI  
 XX  
 XX WPI; 2004-676149/66.  
 DR  
 XX  
 XX Treating patient suffering from debilitating or life threatening disease  
 PT such as malignancy, by administering autophilic antibody capable of  
 PT binding to cellular receptors to promote target cell apoptosis, to  
 PT patient.  
 PT  
 XX  
 PS Example 2; SEQ ID NO 2; 9pp; English.  
 XX  
 XX The invention relates to treating a patient suffering from a debilitating  
 CC or life threatening disease, involving administering at least one  
 CC autophilic antibody to the patient in an amount effective to alleviate  
 CC symptoms of the disease. An autophilic antibody is a non-covalent,  
 CC dimerising antibody formed by the attachment of a peptide which induces  
 CC dimerisation. Also included are potentiating apoptosis of selected cells  
 CC in a patient (involving administering to the patient a first autophilic  
 CC antibody-peptide conjugate and a second antibody directed to the  
 CC autophilic peptide itself), producing an autophilic antibody by  
 CC chemical/genetic engineering techniques (where the autophilic antibody  
 CC contains a T15 autophilic peptide sequence appearing as ADR88403 that  
 CC attaches through a tryptophan photoactivation crosslinking to the  
 CC immunoglobulin component of the antibody, or a modified T15 autophilic  
 CC peptide sequence that further potentiates the ability of the modified  
 CC auto antibody to crosslink once bound to a target antigen relative to the  
 CC unmodified auto antibody), formulating an autophilic antibody composition  
 CC so as to reduce or mitigate dimerisation in solution (comprising addition  
 CC of salt concentrations of 0.5M or more, low levels of sodium dodecyl  
 CC sulphate (SDS), various detergents especially those of an anionic nature,  
 CC or modifications of the antibody to decrease its isoelectric point as  
 CC with succinyl anhydride), expressing an increased degree of apoptosis in  
 CC an in vitro assay of an antibody/antigen system (involving employing an  
 CC autophilic conjugate), identifying an autophilic antibody candidate for  
 CC use in humans (involving administering the autophilic antibody to severe  
 CC combined immuno deficiency (SCID) or nude mice having human tumour  
 CC xenografts), determining a peptide sequence for enhanced noncovalent  
 CC autophilic coupling between antibody molecules (involving providing  
 CC several synthetic peptides each having one or more conservative  
 CC substitutions at amino acid positions of a template peptide, and  
 CC comparing self-binding properties of such synthetic peptides relative to  
 CC those of the template peptide) and a conjugate antibody where two or more  
 CC bioactive peptides are conjugated to different sites of an antibody to  
 CC potentiate therapeutic efficacy. The method is useful for treating a  
 CC patient suffering from a debilitating or life threatening disease such as  
 CC malignancy (cancer or tumour), an auto-immune disorder, transplantation  
 CC rejection, Alzheimer's disease, or other neuro-degenerative condition.  
 CC The present sequence is the mouse MTS membrane translocation peptide  
 CC which may be incorporated into the autophilic antibody to facilitate  
 CC entry into the target cell.  
 CC  
 XX Sequence 17 AA;  
 SQ

Query Match 80.9%; Score 55; DB 8; Length 17;  
 Best Local Similarity 85.7%; Pred. No. 0.013;  
 Matches 12; Conservative 1; Mismatches 0; Gaps 0;

Qy 2 KAAVLLPVLAAAP 15  
 : |||||  
 Db 3 EGAVLLPVLAAAP 16

RESULT 34  
 AAY44160  
 ID AAY44160 standard; protein; 12 AA.  
 XX AC  
 XX AAY44160;  
 XX DT  
 XX 01-FEB-2000 (first entry)  
 XX Membrane-translocating peptide sequence.  
 XX DE  
 XX Membrane-translocating peptide sequence; MTS; fusion protein; production;  
 KW Schistosoma japonicum; Glutathione S transferase; adenovirus; mammal;  
 KW p53; immune response; hepatitis B virus; surface antigen; canine; feline;  
 KW protease inhibitor; cancer; tumor suppressor; bovine.  
 XX OS  
 XX Synthetic.  
 XX PN WO9949879-A1.  
 XX PD 07-OCT-1999.  
 XX PF 31-MAR-1999; 99WO-US007189.  
 XX PR 31-MAR-1998; 98US-0080083P.  
 XX PR 04-NOV-1998; 98US-00186170.  
 XX PA (UTVA-) UNIV VANDERBILT.  
 XX PI Lin Y, Donahue JP, Rojas M, Tan ZJ;  
 XX WPI; 1999-610819/52.  
 XX DR N-PSDB; AA228749.  
 XX New peptides containing a membrane-translocating sequence used to develop  
 products for use in, e.g. vaccines.  
 XX Claim 1; Page 66; 85pp; English.

This sequence represents a novel membrane-translocating peptide sequence (MTS). The invention relates to the use of the MTS peptides for generating fusion proteins which can be used for the production of polypeptides of interest such as Schistosoma japonicum glutathione S transferase, an adenovirus E3 19K protein or a mammalian p53 protein. Fusions of the peptides can also be used for inducing an immune response in a mammal using e.g. a viral polypeptide such as hepatitis B surface antigen. They can also be used for protecting a subject from an infectious agent using a polypeptide that inhibits reproduction of the infectious agent such as a protease inhibitor. They can also be used for treating cancer using a polypeptide tumor suppressor such as p53 protein or a polypeptide inhibitor of Bcl-2. The methods can be used for treating canine, feline and bovine diseases and also for studying intracellular proteins

Qy 4 AAVLLPVLAAAP 15  
 : |||||  
 Db 1 AAVLLPVLAAAP 12

Query Match 79.4%; Score 54; DB 2; Length 12;  
 Best Local Similarity 100.0%; Pred. No. 0.013;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 AAVLLPVLAAAP 15  
 : |||||  
 Db 1 AAVLLPVLAAAP 12

RESULT 35  
 AAE02980  
 ID AAE02980 standard; peptide; 12 AA.  
 XX AC AAE02980;  
 XX DT  
 XX 10-AUG-2001 (first entry)  
 XX Hydrophobic protein transduction domain #8.  
 XX DE  
 XX Peptide monomer; nuclear localisation sequence; NLS;  
 KW protein transduction domain; PTD; molecule transfer.  
 XX OS  
 XX Unidentified.  
 XX PN WO200138547-A2.  
 XX PD 31-MAY-2001.  
 XX PF 23-NOV-2000; 2000WO-EP011690.  
 XX PR 24-NOV-1999; 99EP-00123423.  
 XX PA (ROSE/) ROSENECKER J.  
 XX PA (RITT/) RITTER W.  
 XX PA (RUDO/) RUDOLPH C M.  
 XX PA (PLAN/) PLANK C.  
 XX PI Rosenecker J, Ritter W, Rudolph CM, Plank C;  
 XX WPI; 2001-367696/38.  
 XX Novel polypeptides comprising at least two monomers which comprise a  
 nuclear localization sequence and protein transduction domain,  
 PT respectively useful for transferring nucleic acid molecules into  
 eukaryotic cells.  
 XX Claim 3; Page 32; 68pp; English.

The present invention relates to a polypeptide comprising at least two peptide monomers, in which each peptide monomer comprises an amino acid sequence which serves as a nuclear localisation sequence (NLS) or an amino acid sequence which serves as protein transduction domain (PTD) in eukaryotic cells. The polypeptide of the invention is used for transferring a molecule into eukaryotic cells. The use of a polypeptide comprising NLS or PTD drastically increases the efficiency of the transfer of attached molecules, preferably negatively charged molecules into the nucleus or cytoplasm of a eukaryotic cell. The present sequence is hydrophobic PTD which include the sequence of transportan

Qy 4 AAVLLPVLAAAP 15  
 : |||||  
 Db 1 AAVLLPVLAAAP 12

Query Match 79.4%; Score 54; DB 4; Length 12;  
 Best Local Similarity 100.0%; Pred. No. 0.013;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 AAVLLPVLAAAP 15  
 : |||||  
 Db 1 AAVLLPVLAAAP 12

RESULT 36  
 AAU00644  
 ID AAU00644 standard; peptide; 12 AA.  
 XX AC AAU00644;  
 XX DT  
 XX 07-SEP-2001 (first entry)  
 XX Human membrane translocating peptide (MTLP) #13.  
 XX Membrane translocating peptide; MTLP; human; intracellular gene delivery;

KW epithelial cell layer; gastrointestinal tract; circulatory system.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200127154-A2.  
 XX  
 PD 19-APR-2001.  
 XX  
 XX 27-SEP-2000; 2000WO-IB001491.  
 PF 27-SEP-1999; 99US-0156246P.  
 XX  
 PR (OMAH/) O'MAHONY D J.  
 PA (LAMB/) LAMBKIN I J.  
 XX  
 XX O'mahony DJ, Lambkin IJ;  
 PI WPI; 2001-300212/31.  
 XX  
 DR N-PSDB; AAS00638.  
 XX  
 XX Compositions for enhancing uptake of e.g. drugs or DNA across a cell  
 PT membrane, comprise membrane translocating peptides having specific amino  
 PT acid sequences or a derivative, fragment, motif, analog or peptidomimetic  
 PT of the peptides.  
 XX  
 PS Claim 2; Page 11; 42pp; English.  
 XX  
 CC The sequence represents a human membrane translocated peptide (MTLP).  
 CC MTLPs and their related fragments, motifs, derivatives and analogues are  
 CC used for enhancing uptake of a pharmaceutically active agent into a cell,  
 CC into or out of an intracellular compartment and across a cell layer (for  
 CC example, an epithelial cell layer lining the gastrointestinal tract),  
 CC either directly or from a pharmaceutically active agent loaded particle,  
 CC into the circulatory system of an animal. This method is useful for  
 CC intracellular gene delivery, as a rapid screening method for the  
 CC identification of MTLPs which retain the functional activity of a full-  
 CC length MTLP, as a cell-based screen for assaying the functional activity  
 CC of a MTLP and characterising the properties of a MTLP, for diagnosis of a  
 CC pathological disorder (by administration of a MTLP-active agent complex  
 CC or MTLP-active particle comprising a diagnostic agent) and for  
 CC preventing or treating a pathological disorder  
 XX  
 SQ Sequence 12 AA;  
 Query Match 79.4%; Score 54; DB 4; Length 12;  
 Best Local Similarity 100.0%; Pred. No. 0.013;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 4 AAVLLPVLAAAP 15  
 Db 1 AAVLLPVLAAAP 12  
 |||||  
 |||||  
 RESULT 37  
 ABG78983  
 ID ABG78983 standard; peptide; 12 AA.  
 XX  
 AC ABG78983;  
 XX  
 DT 15-NOV-2002 (first entry)  
 XX  
 DE Cell penetrating peptide CPp1.  
 XX  
 KW Cell penetrating peptide; cancer; tumour; melanoma; thymoma; lymphoma;  
 KW sarcoma; lung cancer; non-Hodgkin's lymphoma; leukaemia;  
 KW Hodgkin's lymphoma; uterine cancer; cervical cancer; bladder cancer;  
 KW kidney cancer; adenocarcinoma; breast cancer; prostate cancer;  
 KW ovarian cancer; pancreatic cancer; vaccine; dendritic cell;  
 KW tumour infiltrating lymphocyte; TIL; human leukocyte antigen; HLA;  
 KW cytostatic.  
 XX  
 OS Unidentified.  
 XX  
 PN WO200207752-A2.  
 XX  
 PD 31-JAN-2002.

PN WO200264057-A2.  
 XX  
 PD 22-AUG-2002.  
 XX  
 PF 15-FEB-2002; 2002WO-US005212.  
 XX  
 PR 15-FEB-2001; 2001US-0268687P.  
 XX  
 PA (BAYU ) BAYLOR COLLEGE MEDICINE.  
 XX  
 PI Wang R;  
 XX  
 DR WPI; 2002-627577/67.  
 XX  
 PT Novel composition for treating a disease in an animal, comprises an  
 PT immune effector cell and cell penetrating peptide associated with an  
 PT antigen or antibody.  
 XX  
 PS Disclosure; Page 10; 61pp; English.  
 XX  
 CC The invention relates to a composition (I) comprising an immune effector  
 CC cell and a cell penetrating peptide (CPP) associated with an antigen or  
 CC antibody. Also included are (i) a vaccine comprising (I), CPP associated  
 CC with an antigen, and a pharmaceutically acceptable carrier and (2)  
 CC preparing a composition for a disease, by providing (I) and CPP  
 CC associated with an antigen for disease, and introducing the antigen-  
 CC associated CPP to (I), where antigen enters into the cell. The antigens  
 CC are, for example, tumour antigen derived epitopes recognised by tumour  
 CC infiltrating lymphocytes (TIL) of HLA (human leukocyte antigen) class I  
 CC or II. The composition is useful for enhancing immunity in an animal to a  
 CC disease, by administering a mature dendritic cell comprising CPP  
 CC associated with an antigen to disease, to the animal, such that following  
 CC the administration, animal is protected from disease, where the animal  
 CC comprises both CD4+ and CD8+ T cells. It is also useful for treating a  
 CC disease (e.g. cancer, tumour, melanoma, thymoma, lymphoma, sarcoma, lung  
 CC cancer, non-Hodgkin's lymphoma, leukaemia, kidney cancer, adenocarcinoma,  
 CC cancer, cervical cancer, bladder cancer, ovarian cancer and pancreatic cancer).  
 CC breast cancer, prostate cancer, cancer, ovarian cancer and pancreatic cancer).  
 CC The animal is further subjected to a cancer treatment including surgery,  
 CC radiation, chemotherapy or gene therapy. The administration of (I),  
 CC preferably dendritic cell is prior to, subsequent to or concurrent with,  
 CC the cancer treatment. The present sequence is cell penetrating peptide of  
 CC the invention  
 XX  
 SQ Sequence 12 AA;  
 Query Match 79.4%; Score 54; DB 5; Length 12;  
 Best Local Similarity 100.0%; Pred. No. 0.013;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 4 AAVLLPVLAAAP 15  
 Db 1 AAVLLPVLAAAP 12  
 |||||  
 |||||  
 RESULT 38  
 AAU78350  
 ID AAU78350 standard; peptide; 12 AA.  
 XX  
 AC AAU78350;  
 XX  
 DT 18-JUN-2002 (first entry)  
 XX  
 DE Grb2 SH2 domain derived signal peptide.  
 XX  
 KW Membrane translocation signal; signal sequence based peptide I;  
 KW red blood cell vehicle; polypeptide delivery; Grb2; SH2 domain.  
 XX  
 OS Unidentified.  
 XX  
 PN WO200207752-A2.  
 XX  
 PD 31-JAN-2002.



CC of an agent to a vertebrate, the red blood cell comprising the novel  
 CC agent-MTS conjugate and a method of immunisation of an animal with an  
 CC antigen. The method is useful for preparing delivery vehicles,  
 CC particularly a red blood cell, for the intracellular delivery of a  
 CC therapeutic agent to a target site. The method is particularly useful for  
 CC enabling an agent to cross the plasma membrane of a target cell, and for  
 CC selectively releasing the agent-MTS conjugate at a target site to  
 CC facilitate the uptake of the agent by the cells at the target site. The  
 CC preferred MTS sequences are from HIV-1 trans-activating protein (Tat),  
 CC Drosophila Antennapedia homeodomain protein (Antp-HD), Herpes Simplex-1  
 CC virus VP22 protein (HSV-VP22), a signal-sequence-based peptide and a  
 CC Transportan and Amphiphilic model peptide. The sequence presented is the  
 CC Signal-sequence-based peptide III, which is derived from the Grb2 SH2  
 CC (not defined) domain  
 CC  
 XX Sequence 12 AA;  
 SQ

Query Match 79.4%; Score 54; DB 5; Length 12;  
 Best Local Similarity 100.0%; Pred. No. 0.013;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 AAVLLPVLLAAP 15  
 Db 1 AAVLLPVLLAAP 12  
 |||||

RESULT 41  
 ABB81178  
 ID ABB81178 standard; peptide; 12 AA.  
 AC ABB81178;  
 XX  
 DT 25-NOV-2002 (first entry)  
 XX  
 DE Grb2 SH2 domain derived peptide fragment.  
 XX  
 KW Red blood cell; RBC; virus; vertebrate; therapeutic; tissue imaging;  
 KW Grb2; SH2 domain; membrane translocation.  
 XX  
 OS Unidentified.  
 XX  
 PN WO200260416-A1.  
 XX  
 PD 08-AUG-2002.  
 XX  
 PF 01-FEB-2002; 2002WO-GB000437.  
 XX  
 PR 01-FEB-2001; 2001GB-00002561.  
 PR 16-FEB-2001; 2001US-0269528P.  
 XX  
 PA (GENE-) GENDEL LTD.  
 XX  
 PI Mchale AP, Craig R;  
 XX  
 DR WPI; 2002-643355/69.  
 XX  
 PT Delivering agent to target site in vertebrate comprises loading red blood  
 PT cell with virus or virus-like particle comprising agent, sensitizing  
 PT cell, introducing cell into vertebrate and applying energy to release  
 PT virus particle from cell.  
 XX  
 PS Disclosure; Page 55; 87pp; English.  
 XX

CC The invention relates to delivering an agent to a target site in a  
 CC vertebrate. The method involves (a) loading a red blood cell (RBC) with a  
 CC virus or a virus-like particle (I) comprising an agent; (b) sensitizing  
 CC RBC to render it more susceptible to disruption than unsensitized RBC;  
 CC (c) introducing RBC into a vertebrate, and (d) applying energy to release  
 CC (i). Steps (a) and (b) may be performed in any order. (ii) RBC loaded  
 CC with a virus or a virus-like particle comprising a therapeutic agent) is  
 CC useful for the delivery of a therapeutic agent to a target site in a  
 CC vertebrate, or in the preparation of a medicament for delivery of a  
 CC therapeutic agent to a target site in a vertebrate. (ii) is also useful

CC for delivering one or more agents to a vertebrate and for treating or  
 CC preventing a disease. The method is useful for delivering agents such as  
 CC those useful for imaging of tissues in vivo or ex vivo, preferably for  
 CC delivering an agent to a subcellular organelle such as nucleus,  
 CC mitochondria, Golgi or endoplasmic reticulum. The present sequence  
 CC represents a Grb2 SH2 domain derived peptide fragment, used for membrane  
 CC translocation  
 XX  
 XX Sequence 12 AA;  
 SQ

Query Match 79.4%; Score 54; DB 5; Length 12;  
 Best Local Similarity 100.0%; Pred. No. 0.013;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 AAVLLPVLLAAP 15  
 Db 1 AAVLLPVLLAAP 12  
 |||||

RESULT 42  
 AAU77231  
 ID AAU77231 standard; peptide; 12 AA.  
 XX  
 AC AAU77231;  
 XX  
 DT 05-JUN-2002 (first entry)  
 XX  
 DE Synthetic pCDNA3-E7/MTS peptide sequence.  
 XX  
 KW Virucide; cytostatic; vaccine; intercellular transport; antigenic;  
 KW immune response; cytotoxic T lymphocyte; tumour; cancer; pCDNA3-E7/MTS;  
 KW chronic viral infection; veterinary herpesvirus infection; pseudorabies;  
 KW equine herpesvirus; bovine herpesvirus; Marek's disease virus; chicken;  
 KW fowl; animal retroviral disease; rabies.  
 XX  
 OS Synthetic.  
 XX  
 PN WO200209645-A2.  
 XX  
 PD 07-FEB-2002.  
 XX  
 PF 01-AUG-2001; 2001WO-US023966.  
 XX  
 PR 01-AUG-2000; 2000US-0222185P.  
 PR 15-FEB-2001; 2001US-0268575P.  
 PR 04-APR-2001; 2001US-0281004P.  
 XX  
 PA (UYTO ) UNIV JOHNS HOPKINS.  
 XX  
 PI Wu T, Hung C;  
 XX  
 DR WPI; 2002-257367/30.  
 DR N-PSDB; ABK11804.  
 XX  
 PT New nucleic acids encoding fusion polypeptide comprising intercellular  
 PT transport polypeptide linked to antigenic polypeptide, useful as  
 PT therapeutic vaccine for cancer and major chronic viral infections.  
 XX  
 PS Example 1; Page 39; 102pp; English.  
 XX

CC The present invention relates to a new nucleic acid molecule that encodes  
 CC a fusion polypeptide. The fusion protein comprises a first polypeptide  
 CC comprising at least one intercellular transport polypeptide and a second  
 CC polypeptide comprising at least one antigenic polypeptide or peptide. The  
 CC invention also describes an optional linker peptide linking the first and  
 CC second polypeptide. The nucleic acid is useful as a vaccine for enhancing  
 CC immune responses, primarily cytotoxic T lymphocyte responses to specific  
 CC antigens such as tumour or viral antigens. The compositions comprising  
 CC the nucleic acids are especially useful as a therapeutic vaccine for  
 CC cancer and for major chronic viral infections, as well as in the  
 CC treatment of veterinary herpesvirus infections, including equine or  
 CC bovine herpesvirus, Marek's disease virus in chickens and other fowls,  
 CC animal retroviral diseases, pseudorabies and rabies. The present amino

CC acid sequence represents the peptide used in the methods of the invention  
CC for the generation of pcDNA3-E7/MTS expression vector

SQ Sequence 12 AA;

Query Match 79.4%; Score 54; DB 5; Length 12;  
Best Local Similarity 100.0%; Pred. No. 0.013;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 AAVLLPVLLAAP 15  
Db 1 AAVLLPVLLAAP 12

RESULT 43

AAE32065  
ID AAE32065 standard; peptide; 12 AA.

XX AC AAE32065;

XX DT 24-MAR-2003 (first entry)

XX DE Transport peptide used in the invention.

XX KW Drug delivery construct; axon growth; nerve injury; ischaemic damage;  
XX KW stroke injury; gene therapy; neuroleptic; neuroprotective.

XX OS Unidentified.

XX PN WO200283179-A2.

XX PD 24-OCT-2002.

XX PF 08-APR-2002; 2002WO-CA000480.

XX PR 12-APR-2001; 2001CA-02342970.

XX PR 13-NOV-2001; 2001CA-02362004.

XX PR 15-JAN-2002; 2002CA-02367636.

XX PA (BIOA-) BIOAXONE THERAPEUTIQUE INC.

XX PI McKerracher L;

XX DR WPI; 2003-092963/08.

XX PT New drug delivery construct comprising a transport and active agent

XX PT region, useful for the manufacture of a pharmaceutical composition for

XX PT treating nerve injury.

XX PS Disclosure; Page 52; 188pp; English.

XX CC The invention relates to a new drug delivery construct comprises at least  
XX CC one transport agent region and an active agent region. The transport  
XX CC agent region is able to facilitate the uptake of the active agent region  
XX CC into a cell. The active agent region is an active therapeutic agent  
XX CC region able to facilitate axon growth and an analogue. The drug delivery  
XX CC construct is useful for suppressing the inhibition of neuronal axon  
XX CC growth, facilitating axon growth, treating nerve injury, treating  
XX CC ischaemic damage related to stroke injury. The drug delivery construct  
XX CC and the drug conjugate are useful for the manufacture of a pharmaceutical  
XX CC composition for treating nerve injury. The invention is useful in gene  
XX CC therapy. The present sequence is transport peptide used in the invention

XX SQ Sequence 12 AA;

Query Match 79.4%; Score 54; DB 6; Length 12;  
Best Local Similarity 100.0%; Pred. No. 0.013;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 AAVLLPVLLAAP 15  
Db 1 AAVLLPVLLAAP 12

RESULT 45

ADC22455

ID ADC22455 standard; peptide; 12 AA.

XX

RESULT 44

ADB88787

ID ADB88787 standard; peptide; 12 AA.

XX AC ADB88787;

XX DT 04-DEC-2003 (first entry)

XX DE Membrane translocating peptide #13.

XX KW Peyer's patch cell; non-Peyer's patch cell; transcription factor;  
XX KW upregulated protein; antigen; vaccine delivery; M cell;  
XX KW membrane translocating peptide.

XX OS Unidentified.

XX PN WO2003004646-A2.

XX PD 16-JAN-2003.

XX PF 04-APR-2002; 2002WO-1B003866.

XX PR 04-APR-2001; 2001US-0281387P.

XX PR 02-JUL-2001; 2001US-0302591P.

XX PA (OMAH/) O'MAHONY D J.

XX PI O'mahony DJ, Byrne D, Brayden D;

XX DR WPI; 2003-229409/22.

XX PT Increasing the levels of a protein in a Peyer's patch cell, useful for  
XX PT targeted vaccine or drug delivery, comprises delivering to the Peyer's  
XX PT patch cell a transcription factor or an activator of a transcription  
XX PT factor.

XX PS Example 6; Page 51; 147pp; English.

XX CC The invention relates to a novel method for increasing the levels of a  
XX CC protein in a Peyer's patch cell. The method comprises delivering to the  
XX CC cell a nucleic acid coding for a protein, the level of which or its mRNA  
XX CC is greater than in a non-Peyer's patch cell. The preferred protein of the  
XX CC invention is a transcription factor or a protein that activates a  
XX CC transcription factor selected from Jun-B, c-jun related TF, Jun-D, STAT 3  
XX CC -signal transducer and activator of transcription 3, Nfkapppabgr; Tf p105  
XX CC subunit, S-myc proto-oncogene, myc related, Nm23-M2, nucleoside  
XX CC diphosphate kinase B, metastasis reducing protein, and C-est-I proto-  
XX CC oncogene, and p54. The preferred upregulated protein of the invention is  
XX CC selected from clusterin, T-cell surface glycoprotein CD5 precursor, HSP  
XX CC 84, and Ca2+ dependent phospholipase A2 precursor and the mRNA is for a  
XX CC protein selected from the group. The method is useful for increasing or  
XX CC decreasing the level of a protein in a Peyer's patch cell, particularly  
XX CC in increasing antigen or vaccine delivery to M cells. The method may also  
XX CC be used to enhance transport of a drug through the gastrointestinal tract  
XX CC (GIT). This sequence represents a membrane translocating peptide of the  
XX CC invention.

XX SQ Sequence 12 AA;

Query Match 79.4%; Score 54; DB 7; Length 12;  
Best Local Similarity 100.0%; Pred. No. 0.013;  
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 AAVLLPVLLAAP 15  
Db 1 AAVLLPVLLAAP 12

AC ADC2455;  
 XX  
 DT 18-DEC-2003 (first entry)  
 XX  
 DE Protein-derived transport peptide SEQ ID NO:304.  
 DE  
 XX recombinant fusion protein; fusion protein; binding; detection;  
 KW localisation domain; binding domain;  
 KW subcellular compartment localisation.  
 XX  
 OS Synthetic.  
 XX  
 XX WO2003012068-A2.  
 PN  
 XX 13-FEB-2003.  
 PD  
 XX  
 PF 01-AUG-2002; 2002WO-US024572.  
 XX  
 PR 01-AUG-2001; 2001US-0309395P.  
 PR  
 XX 13-DEC-2001; 2001US-0341589P.  
 XX  
 PA (CELL-) CELLONICS INC.  
 XX  
 XX Bright G, Premkumar DR, Chen Y;  
 XX  
 XX WPI; 2003-248174/24.  
 DR  
 XX  
 XX New recombinant fusion protein comprising detection and first  
 PT localization domains and a binding domain for the molecule of interest,  
 PT useful for detecting binding of a molecule of interest.  
 XX  
 XX Disclosure; SEQ ID NO 304; 101pp; English.  
 XX  
 CC The present invention describes a recombinant fusion protein (I) for  
 CC detecting binding of a molecule of interest. (I) comprises: (a) a  
 CC detection domain; (b) a first localisation domain; and (c) a binding  
 CC domain for the molecule of interest. The detection domain, the first  
 CC localisation domain and the binding domain for the molecule of interest  
 CC constituting the recombinant fusion protein for detecting binding of a  
 CC molecule of interest are operably linked. The binding domain for the  
 CC molecule of interest is separated from the first localisation domain by 0  
 CC -20 amino acid residues. The first localisation domain and the binding  
 CC domain for the molecule of interest both do not occur in a single non-  
 CC recombinant protein with the same spacing as in the recombinant fusion  
 CC protein for detecting binding of a molecule of interest. Also described:  
 CC (1) a recombinant nucleic acid encoding the recombinant fusion protein;  
 CC (2) a recombinant expression vector comprising the nucleic acid control;  
 CC sequences operably linked to the recombinant nucleic acid molecule; (3) a  
 CC genetically engineered host cell transfected with the recombinant  
 CC expression vector; (4) a kit for detecting binding of the molecule of  
 CC interest; and (5) a method for identifying compounds that alter the  
 CC binding of the molecule of interest. The recombinant fusion protein is  
 CC useful for detecting binding of a molecule of interest. The recombinant  
 CC fusion protein eliminates the need to construct two or more chimeric  
 CC proteins and enables the monitoring of biochemical events in live, intact  
 CC or fixed cells. The present sequence is used in the exemplification of  
 CC the present invention.  
 XX  
 XX Sequence 12 AA;  
 SQ  
 Query Match 79.4%; Score 54; DB 7; Length 12;  
 Best Local Similarity 100.0%; Pred. No. 0.013; 0; Gaps 0;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 4 AAVLLPVLAAAP 15  
 Db 1 AAVLLPVLAAAP 12  
 |||||  
 |||||  
 RESULT 46  
 ADG28018  
 ID ADG28018 standard; peptide; 12 AA.  
 XX

AC ADG28018;  
 XX  
 DT 26-FEB-2004 (first entry)  
 XX  
 DE Kaposi FGF signal membrane fusion sequence seq id 19.  
 DE  
 XX fusion protein; cold shock domain; membrane translocation sequence; CspA;  
 KW CspB, CspC; CspD; rpi S1 binding domain; eukaryotic Y-box protein;  
 KW DNA binding protein B; DBPB; DBPA; EFE-1; mRNP3; mRNP4; FRG Y1;  
 KW nuclease-sensitive element binding protein 1; NSEP 1;  
 KW DNA condensation domain; DNA binding domain; SPKR;  
 KW nuclear localisation sequence; NLS; protein purification tagged sequence;  
 KW gene delivery; kaposi's sarcoma-associated herpesvirus;  
 KW FGF signal sequence; membrane fusion sequence.  
 XX  
 OS Human herpesvirus 8.  
 XX  
 XX US2003211590-A1.  
 PN  
 XX 13-NOV-2003.  
 PD  
 XX 13-MAY-2002; 2002US-00144549.  
 PF  
 XX 13-MAY-2002; 2002US-00144549.  
 PR  
 XX (HWUP/) HWU P L.  
 XX  
 PA Hwu PL;  
 XX  
 XX WPI; 2003-901590/82.  
 DR  
 XX  
 XX New fusion protein comprising a cold shock domain, and a membrane  
 PT translocation sequence, useful for delivering DNAs and RNAs to in vivo  
 PT cells for gene delivery.  
 XX  
 XX Claim 9; SEQ ID NO 19; 24pp; English.  
 XX  
 CC The invention describes a fusion protein for delivery of a desired  
 CC molecule into cells or nuclei, comprising a cold shock domain, its  
 CC homologue and functional derivative, and a membrane translocation  
 CC sequence or its functional equivalent peptides and/or derivatives. The  
 CC fusion protein comprises a cold shock domain that is selected from CspA,  
 CC CspB, CspC, CspD, rpi S1 binding domain, eukaryotic Y-box proteins, DNA  
 CC binding protein B (DBPB), DBPA, EFE-1, mRNP3, mRNP4, FRG Y1 and nuclease-  
 CC sensitive element binding protein 1 (NSEP 1). The functional equivalent  
 CC derivative of cold shock protein is modified by inserting into the cold  
 CC shock domain with a DNA condensation domain or a DNA binding domain. The  
 CC DNA condensation or binding domain is selected from DNA condensation  
 CC domain (SPKR) 3-4 and the positive charge nuclear localisation sequences  
 CC (NLSs). The membrane transduction sequence is protein transduction domain  
 CC (PTD) or membrane fusion sequence. The fusion protein further comprises a  
 CC protein purification tagged sequence selected from HA, GST, and His6 tag.  
 CC The fusion protein is useful for delivering DNAs and RNAs to in vivo  
 CC cells for gene delivery, or for delivering nucleic acids to an embryo or  
 CC to a living animal for the production of transgenic animal. This is the  
 CC amino acid sequence of a membrane fusion sequence derived from Kaposi's  
 CC sarcoma-associated herpesvirus FGF signal sequence.  
 XX  
 XX Sequence 12 AA;  
 SQ  
 Query Match 79.4%; Score 54; DB 7; Length 12;  
 Best Local Similarity 100.0%; Pred. No. 0.013; 0; Gaps 0;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 4 AAVLLPVLAAAP 15  
 Db 1 AAVLLPVLAAAP 12  
 |||||  
 |||||  
 RESULT 47  
 ADH76185  
 ID ADH76185 standard; peptide; 12 AA.  
 XX



AC ADH76185;  
 XX 22-APR-2004 (first entry)  
 XX Transduction domain peptide of the invention SEQ ID NO:286.  
 XX heat shock protein 20; cytosolic; antiarteriosclerotic; vasotropic;  
 KW anti-anginal; cerebroprotective; antiarrhythmic; antiasthmatic;  
 KW gynaecological; hypotensive; antimigraine; tocolytic; relaxant; HSP;  
 KW smooth muscle cell; smooth muscle cell; transduction domain.  
 XX Synthetic.  
 OS  
 XX WO2003018758-A2.  
 PN  
 XX 06-MAR-2003.  
 PD  
 XX 23-AUG-2002; 2002WO-US026918.  
 PF  
 XX 23-AUG-2001; 2001US-0314535P.  
 PR  
 XX (UYAR-) UNIV ARIZONA.  
 PA (BROP/) BROPHY C.  
 PA (KONA/) KOMALAVILAS P.  
 PA (PANI/) PANITCH A.  
 PA (SEAL/) SEAL B.  
 PA (LOKE/) LOKESH J.  
 XX Brophy C, Komalavilas P, Panitch A, Seal B, Lokesh J;  
 PI WPI; 2003-393248/37.  
 XX New heat shock protein 20-derived polypeptides, useful for inhibiting,  
 PT treating or preventing smooth muscle cell vasospasm or a disorder such as  
 PT intimal hyperplasia, restenosis, atherosclerosis or smooth muscle cell  
 PT tumors.  
 XX Claim 29; SEQ ID NO 286; 194pp; English.  
 PS The invention relates to a novel polypeptide comprising a heat shock  
 XX protein 20-derived polypeptide. A polypeptide of the invention has  
 CC cytosolic, antiarteriosclerotic, vasotropic, anti-anginal,  
 CC cerebroprotective, antiarrhythmic, antiasthmatic, gynaecological,  
 CC hypotensive, antimigraine, tocolytic, and relaxant activity, and may act  
 CC as a HSP agonist or antagonist. The polypeptides, heat shock protein  
 CC (HSP) 20, and methods are useful for treating or preventing a disorder,  
 CC e.g. intimal hyperplasia, stenosis, restenosis, transplant vasculopathy,  
 CC atherosclerosis, smooth muscle cell tumors such as leiomyosarcoma, or  
 CC vasospasm, which is associated with angina, coronary vasospasm,  
 CC Prinzmetal's angina, coronary ischaemia, stroke, bradycardia,  
 CC hypertension, pulmonary (lung) hypertension, asthma (bronchospasm),  
 CC toxemia of pregnancy, pre-term labour, pre-eclampsia/eclampsia, Raynaud's  
 CC disease, Raynaud's phenomenon, haemolytic-uremia, non-occlusive  
 CC mesenteric ischaemia, anal fissure, achalasia, impotence, migraine, or  
 CC ischaemic muscle injury associated with smooth muscle spasm. The  
 CC polypeptides and HSP 20 are also useful for inhibiting smooth muscle cell  
 CC proliferation and/or migration. The present sequence represents a peptide  
 CC used in a polypeptide of the invention.  
 XX Sequence 12 AA;  
 SQ  
 Query Match 79.4%; Score 54; DB 7; Length 12;  
 Best Local Similarity 100.0%; Pred. No. 0.013;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 4 AAVLLPVLLAAP 15  
 Db 1 AAVLLPVLLAAP 12  
 RESULT 48  
 ADL88654  
 ID ADL88654 standard; peptide; 12 AA.

XX ADL88654;  
 AC 20-MAY-2004 (first entry)  
 DT  
 XX MPS (Kaposi FGF signal sequence) membrane fusion sequence peptide 2.  
 DE fusion protein; cold shock domain; membrane translocation; gene therapy;  
 KW transgenic; membrane fusion; MPS; Kaposi FGF signal.  
 KW Unidentified.  
 OS  
 XX JP2004035409-A.  
 PN  
 XX 05-FEB-2004.  
 PD  
 XX 15-MAY-2002; 2002JP-00140441.  
 PF  
 XX 13-MAY-2002; 2002US-00144549.  
 PR  
 XX (GENE-) GENESHUTTLE BIOPHARM INC.  
 PA  
 XX Hwu PL;  
 PI WPI; 2003-901590/82.  
 XX New fusion protein comprising a cold shock domain, and a membrane  
 PT translocation sequence, useful for delivering DNAs and RNAs to in vivo  
 PT cells for gene delivery.  
 PT Claim 9; SEQ ID NO 17; 53pp; Japanese.  
 XX The invention relates to a novel fusion protein for delivery of a desired  
 CC molecule into cells or nuclei comprising a cold shock domain, its  
 CC homologue and functional derivative and a membrane translocation sequence  
 CC or its functionally equivalent peptides and/or derivatives. The fusion  
 CC protein of the invention may be useful for delivering DNAs and RNAs to in  
 CC vivo cells for gene therapy or for delivering nucleic acids to an embryo  
 CC or to a living animal for the production of transgenic animals. The  
 CC current sequence is that of a membrane fusion sequence peptide of the  
 CC invention.  
 XX Sequence 12 AA;  
 SQ  
 Query Match 79.4%; Score 54; DB 7; Length 12;  
 Best Local Similarity 100.0%; Pred. No. 0.013;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 4 AAVLLPVLLAAP 15  
 Db 1 AAVLLPVLLAAP 12  
 RESULT 49  
 ADN60178  
 ID ADN60178 standard; peptide; 12 AA.  
 XX  
 AC ADN60178;  
 XX 01-JUL-2004 (first entry)  
 DT  
 XX Novel recombinase-related Kaposi growth factor peptide SeqID18.  
 DE fusion protein; site-specific DNA recombinase domain;  
 KW nuclear localisation signal; NLS; gene alteration; cell culture;  
 KW cellular uptake; functional biopolymer; Kaposi fibroblast growth factor.  
 KW Unidentified.  
 OS  
 XX WO2003076561-A2.  
 PN  
 XX 18-SEP-2003.  
 PD  
 XX

PF 06-MAR-2003; 2003WO-EP002280.  
 XX  
 PR 09-MAR-2002; 2002EP-00005468.  
 PR 13-MAR-2002; 2002US-0363757P.  
 XX  
 XX (ARTE-) ARTEMIS PHARM GMBH.  
 XX PA  
 XX Edenhofer FOS, Peitz M, Pfannkuche K, Rajewski K;  
 XX WPI; 2003-767415/72.  
 DR  
 XX New fusion protein comprising a site-specific DNA recombinase domain and  
 PT a domain containing a modified nuclear localization signal, useful for  
 PT preparing an agent for inducing target gene alterations in living  
 PT organisms.  
 XX  
 XX Claim 7; SEQ ID NO 18; 54pp; English.  
 XX  
 XX This invention relates to a novel fusion protein comprising a site-  
 CC specific DNA recombinase domain and a domain containing a modified  
 CC nuclear localisation signal (NLS) of type one having 5-10 amino acid  
 CC residues and containing at least 5 basic amino acid residues and no Pro  
 CC residue. The fusion protein is useful for preparing an agent for inducing  
 CC target gene alterations in living organisms or in cell cultures, where  
 CC the living organisms or cells of the cell cultures carry at least one or  
 CC more recognition sites for the site-specific DNA recombinase integrated  
 CC in its genome. The modified NLS is useful for enhancing cellular uptake  
 CC of functional biopolymers in living organisms or cell cultures. The  
 CC present sequence is that of a Kaposi fibroblast growth factor peptide  
 CC which is related to the novel recombinase fusion proteins of the  
 CC invention.  
 XX  
 XX Sequence 12 AA;  
 SQ

Query Match 79.4%; Score 54; DB 7; Length 12;  
 Best Local Similarity 100.0%; Pred. No. 0.013;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 AAVLLPVLLAAP 15  
 DB 1 AAVLLPVLLAAP 12  
 |||||  
 |||||

RESULT 50  
 ADG12916  
 ID ADG12916 standard; peptide; 12 AA.  
 XX  
 XX AC ADG12916;  
 XX  
 XX DT 26-FEB-2004 (first entry)  
 XX  
 XX DE Cytoplasmic transduction peptide WTS.  
 XX  
 XX KW cytoplasmic transduction peptide; cell membrane transduction potential;  
 KW cell membrane transduction; cytotoxic T lymphocyte; drug delivery system.  
 XX  
 XX OS Synthetic.  
 XX  
 XX PN WO2003097671-A1.  
 XX  
 XX PD 27-NOV-2003.  
 XX  
 XX PF 28-MAR-2003; 2003WO-KR000630.  
 XX  
 XX PR 29-MAR-2002; 2002KR-00017546.  
 XX  
 XX PA (CREA-) CREAGENE INC.  
 XX  
 XX PI Kim D, Oh H, Kim C, Kim J, Jeon C, Kim K, Bae Y, Choi I;  
 XX WPI; 2004-061978/06.  
 DR N-PSDB; ADG12915.  
 DR  
 XX

PT New cytoplasmic transduction peptide having a characterized cell membrane  
 PT in the transduction potential, useful for inducing cytotoxic T  
 PT lymphocytes and a drug delivery system targeting cytoplasm.  
 XX  
 XX Example 1; Page 21; 77pp; English.  
 XX  
 CC The present invention describes a cytoplasmic transduction peptide which  
 CC has cell membrane transduction potential, where a cell is treated with  
 CC the cytoplasmic transduction peptide for a period of time and then  
 CC treated with a protease, a cell membrane transduction by the cytoplasmic  
 CC transduction peptide continues to occur, and the cytoplasmic transduction  
 CC remains in the cytoplasm of the cell. Also described: (1) a nucleic acid  
 CC molecule encoding the cytoplasmic transduction peptide; (2) a cytoplasmic  
 CC transduction system comprising the cytoplasmic transduction peptide and a  
 CC biologically active molecule covalently linked to the peptide; and (3) a  
 CC method for delivering the biologically active molecule into the cytoplasm  
 CC of a cell. The cytoplasmic transduction peptide is useful in inducing  
 CC cytotoxic T lymphocytes and a drug delivery system targeting cytoplasm.  
 CC The present sequence is used in the exemplification of the present  
 CC invention.

XX Sequence 12 AA;

Query Match 79.4%; Score 54; DB 8; Length 12;  
 Best Local Similarity 100.0%; Pred. No. 0.013;  
 Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 AAVLLPVLLAAP 15  
 DB 1 AAVLLPVLLAAP 12  
 |||||  
 |||||

Search completed: June 2, 2005, 01:35:25  
 Job time : 130.688 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 2, 2005, 01:31:46 ; Search time 87.6562 Seconds  
(without alignments)  
59.154 Million cell updates/sec

Title: US-10-764-235-2

Perfect score: 68

Sequence: 1 KKAAYLLPVLAAAP 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1465611 seqs, 345679903 residues

Total number of hits satisfying chosen parameters: 1465611

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 750 summaries

Database : Published Applications AA:\*

1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US05\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep.\*  
17: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep.\*  
19: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
20: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	68	100.0	15	14	US-10-126-845-2
2	68	100.0	15	14	US-10-126-845-3
3	68	100.0	15	14	US-10-126-845-60
4	68	100.0	15	14	US-10-126-845-61
5	68	100.0	15	15	US-10-136-187-2
6	68	100.0	15	15	US-10-136-187-6
7	68	100.0	15	15	US-10-136-187-12
8	68	100.0	15	15	US-10-136-187-13
9	68	100.0	15	15	US-10-136-187-33
10	68	100.0	15	15	US-10-136-187-34
11	68	100.0	15	15	US-10-136-187-35
12	68	100.0	15	15	US-10-136-187-36
13	68	100.0	15	15	US-10-136-187-37
					Sequence 2, Appli
					Sequence 3, Appli
					Sequence 60, Appli
					Sequence 61, Appli
					Sequence 2, Appli
					Sequence 12, Appli
					Sequence 13, Appli
					Sequence 33, Appli
					Sequence 34, Appli
					Sequence 35, Appli
					Sequence 36, Appli
					Sequence 37, Appli

14	68	100.0	15	15	US-10-136-187-38	Sequence 38, Appli
15	68	100.0	15	15	US-10-136-187-39	Sequence 39, Appli
16	68	100.0	15	16	US-10-764-235-2	Sequence 2, Appli
17	68	100.0	15	17	US-10-850-873-2	Sequence 2, Appli
18	68	100.0	15	17	US-10-850-873-6	Sequence 6, Appli
19	68	100.0	15	17	US-10-850-873-12	Sequence 12, Appli
20	68	100.0	15	17	US-10-850-873-13	Sequence 13, Appli
21	68	100.0	15	17	US-10-850-873-33	Sequence 33, Appli
22	68	100.0	15	17	US-10-850-873-34	Sequence 34, Appli
23	68	100.0	15	17	US-10-850-873-35	Sequence 35, Appli
24	68	100.0	15	17	US-10-850-873-36	Sequence 36, Appli
25	68	100.0	15	17	US-10-850-873-37	Sequence 37, Appli
26	68	100.0	15	17	US-10-850-873-38	Sequence 38, Appli
27	68	100.0	15	17	US-10-850-873-39	Sequence 39, Appli
28	68	100.0	15	17	US-10-955-656-2	Sequence 2, Appli
29	68	100.0	15	17	US-10-955-656-3	Sequence 3, Appli
30	68	100.0	15	17	US-10-955-656-60	Sequence 60, Appli
31	68	100.0	15	17	US-10-955-656-61	Sequence 61, Appli
32	68	100.0	16	14	US-10-126-845-48	Sequence 48, Appli
33	68	100.0	16	15	US-10-136-187-3	Sequence 3, Appli
34	68	100.0	16	15	US-10-136-187-7	Sequence 7, Appli
35	68	100.0	16	15	US-10-136-187-28	Sequence 28, Appli
36	68	100.0	16	15	US-10-136-187-32	Sequence 32, Appli
37	68	100.0	16	15	US-10-116-275-91	Sequence 91, Appli
38	68	100.0	16	16	US-10-764-235-3	Sequence 3, Appli
39	68	100.0	16	16	US-10-764-235-48	Sequence 48, Appli
40	68	100.0	16	17	US-10-850-873-3	Sequence 3, Appli
41	68	100.0	16	17	US-10-850-873-7	Sequence 7, Appli
42	68	100.0	16	17	US-10-850-873-28	Sequence 28, Appli
43	68	100.0	16	17	US-10-850-873-32	Sequence 32, Appli
44	68	100.0	16	17	US-10-955-656-48	Sequence 48, Appli
45	68	100.0	17	15	US-10-136-187-30	Sequence 30, Appli
46	68	100.0	17	17	US-10-850-873-30	Sequence 30, Appli
47	68	100.0	19	14	US-10-126-845-4	Sequence 4, Appli
48	68	100.0	19	14	US-10-126-845-62	Sequence 62, Appli
49	68	100.0	19	14	US-10-126-845-108	Sequence 108, App
50	68	100.0	19	15	US-10-116-275-92	Sequence 92, Appli
51	68	100.0	19	16	US-10-764-235-4	Sequence 4, Appli
52	68	100.0	19	17	US-10-955-656-4	Sequence 4, Appli
53	68	100.0	19	17	US-10-955-656-62	Sequence 62, Appli
54	68	100.0	19	17	US-10-955-656-108	Sequence 108, App
55	68	100.0	20	14	US-10-126-845-106	Sequence 106, App
56	68	100.0	20	14	US-10-126-845-107	Sequence 107, App
57	68	100.0	20	15	US-10-136-187-5	Sequence 5, Appli
58	68	100.0	20	15	US-10-136-187-9	Sequence 9, Appli
59	68	100.0	20	17	US-10-850-873-5	Sequence 5, Appli
60	68	100.0	20	17	US-10-850-873-9	Sequence 9, Appli
61	68	100.0	20	17	US-10-955-656-106	Sequence 106, App
62	68	100.0	20	17	US-10-955-656-107	Sequence 107, App
63	68	100.0	21	14	US-10-126-845-109	Sequence 109, App
64	68	100.0	21	14	US-10-126-845-111	Sequence 111, App
65	68	100.0	21	15	US-10-116-275-90	Sequence 90, Appli
66	68	100.0	21	17	US-10-955-656-109	Sequence 109, App
67	68	100.0	21	17	US-10-955-656-111	Sequence 111, App
68	64	94.1	16	14	US-10-126-845-5	Sequence 5, Appli
69	64	94.1	16	14	US-10-126-845-63	Sequence 63, Appli
70	64	94.1	16	15	US-10-116-275-93	Sequence 93, Appli
71	64	94.1	16	15	US-10-764-235-5	Sequence 5, Appli
72	64	94.1	16	17	US-10-955-656-5	Sequence 5, Appli
73	64	94.1	16	17	US-10-955-656-63	Sequence 63, Appli
74	63	92.6	15	15	US-10-136-187-4	Sequence 4, Appli
75	63	92.6	15	15	US-10-136-187-8	Sequence 8, Appli
76	63	92.6	15	15	US-10-136-187-29	Sequence 29, Appli
77	63	92.6	15	17	US-10-850-873-4	Sequence 4, Appli
78	63	92.6	15	17	US-10-850-873-8	Sequence 8, Appli
79	63	92.6	15	17	US-10-850-873-29	Sequence 29, Appli
80	61	89.7	15	15	US-10-136-187-31	Sequence 31, Appli
81	61	89.7	15	15	US-10-136-187-40	Sequence 40, Appli
82	61	89.7	15	17	US-10-850-873-31	Sequence 31, Appli
83	61	89.7	15	17	US-10-850-873-40	Sequence 40, Appli
84	57	83.8	18	16	US-10-784-309A-1	Sequence 1, Appli
85	57	83.8	18	16	US-10-784-309A-2	Sequence 2, Appli
86	55	80.9	17	16	US-10-652-864-2	Sequence 2, Appli

87	55	80.9	17	17	US-10-795-081A-8	Sequence 8, Appli	160	47	69.1	12	16	US-10-764-235-16	Sequence 16, Appl
88	54	79.4	12	9	US-09-789-836-12	Sequence 12, Appl	161	47	69.1	12	16	US-10-764-235-59	Sequence 59, Appl
89	54	79.4	12	9	US-09-785-802A-10	Sequence 10, Appl	162	47	69.1	12	17	US-10-955-656-16	Sequence 16, Appl
90	54	79.4	12	13	US-09-997-465B-4	Sequence 4, Appli	163	47	69.1	12	17	US-10-955-656-74	Sequence 74, Appl
91	54	79.4	12	13	US-10-116-288-1	Sequence 1, Appli	164	47	69.1	13	14	US-10-126-845-6	Sequence 6, Appli
92	54	79.4	12	14	US-10-226-956-286	Sequence 286, App	165	47	69.1	13	14	US-10-126-845-64	Sequence 64, Appl
93	54	79.4	12	14	US-10-077-555-1	Sequence 1, Appli	166	47	69.1	13	15	US-10-116-275-94	Sequence 94, Appl
94	54	79.4	12	14	US-10-211-088-304	Sequence 304, App	167	47	69.1	13	15	US-10-764-235-6	Sequence 6, Appli
95	54	79.4	12	14	US-10-156-570A-27	Sequence 27, Appl	168	47	69.1	13	17	US-10-955-656-6	Sequence 6, Appli
96	54	79.4	12	14	US-10-126-845-1	Sequence 1, Appli	169	47	69.1	13	17	US-10-955-656-64	Sequence 64, Appl
97	54	79.4	12	14	US-10-126-845-14	Sequence 14, Appl	170	47	69.1	13	17	US-10-126-845-23	Sequence 23, Appl
98	54	79.4	12	14	US-10-126-845-72	Sequence 72, Appl	171	47	69.1	17	14	US-10-126-845-81	Sequence 81, Appl
99	54	79.4	12	15	US-10-136-187-1	Sequence 1, Appli	172	47	69.1	17	15	US-10-116-275-111	Sequence 111, App
100	54	79.4	12	15	US-10-116-275-102	Sequence 102, App	173	47	69.1	17	16	US-10-764-235-33	Sequence 23, Appl
101	54	79.4	12	15	US-10-144-549-19	Sequence 19, Appl	174	47	69.1	17	17	US-10-955-656-23	Sequence 23, Appl
102	54	79.4	12	15	US-10-361-208-473	Sequence 473, App	175	47	69.1	17	17	US-10-955-656-81	Sequence 81, Appl
103	54	79.4	12	15	US-10-232-410-2	Sequence 2, Appli	176	46	67.6	10	13	US-10-116-288-8	Sequence 8, Appli
104	54	79.4	12	15	US-10-416-285-1	Sequence 1, Appli	177	46	67.6	10	14	US-10-126-845-20	Sequence 20, Appl
105	54	79.4	12	16	US-10-764-235-1	Sequence 1, Appli	178	46	67.6	10	14	US-10-126-845-78	Sequence 78, Appl
106	54	79.4	12	16	US-10-764-235-14	Sequence 14, Appl	179	46	67.6	10	15	US-10-116-275-108	Sequence 108, App
107	54	79.4	12	16	US-10-751-380-8	Sequence 8, Appli	180	46	67.6	10	16	US-10-764-235-20	Sequence 20, Appl
108	54	79.4	12	17	US-10-706-738-1	Sequence 1, Appli	181	46	67.6	10	17	US-10-634-645-8	Sequence 8, Appli
109	54	79.4	12	17	US-10-850-873-1	Sequence 1, Appli	182	46	67.6	10	17	US-10-955-656-20	Sequence 20, Appl
110	54	79.4	12	17	US-10-634-645-1	Sequence 9, Appli	183	46	67.6	10	17	US-10-955-656-78	Sequence 78, Appl
111	54	79.4	12	17	US-10-795-081A-9	Sequence 9, Appli	184	46	67.6	11	17	US-10-916-185-17	Sequence 17, Appl
112	54	79.4	12	17	US-10-823-259-38	Sequence 38, Appl	185	46	67.6	12	14	US-10-126-845-11	Sequence 11, Appl
113	54	79.4	12	17	US-10-823-254-38	Sequence 38, Appl	186	46	67.6	12	14	US-10-126-845-59	Sequence 59, Appl
114	54	79.4	12	17	US-10-902-959-51	Sequence 51, Appl	187	46	67.6	12	15	US-10-116-275-99	Sequence 99, Appl
115	54	79.4	12	17	US-10-955-656-1	Sequence 1, Appli	188	46	67.6	12	16	US-10-764-235-11	Sequence 11, Appl
116	54	79.4	12	17	US-10-955-656-14	Sequence 14, Appl	189	46	67.6	12	17	US-10-955-656-11	Sequence 11, Appl
117	54	79.4	12	17	US-10-955-656-72	Sequence 72, Appl	190	46	67.6	12	17	US-10-955-656-69	Sequence 69, Appl
118	54	79.4	13	14	US-10-013-815-20	Sequence 20, Appl	191	45	66.2	356	16	US-10-437-963-115954	Sequence 115954,
119	54	79.4	386	10	US-09-948-193-1	Sequence 1, Appli	192	44	64.7	17	14	US-10-126-845-104	Sequence 104, App
120	53	77.9	14	14	US-10-126-845-7	Sequence 7, Appli	193	44	64.7	17	17	US-10-955-656-104	Sequence 104, App
121	53	77.9	14	14	US-10-126-845-65	Sequence 65, Appl	194	43	63.2	10	13	US-10-116-288-4	Sequence 4, Appli
122	53	77.9	14	15	US-10-116-275-95	Sequence 95, Appl	195	43	63.2	10	17	US-10-634-645-4	Sequence 4, Appli
123	53	77.9	14	16	US-10-764-235-7	Sequence 7, Appli	196	43	63.2	186	15	US-10-282-122A-69222	Sequence 69222, A
124	53	77.9	14	16	US-10-764-235-58	Sequence 58, Appl	197	42	61.8	9	13	US-10-116-288-7	Sequence 7, Appli
125	53	77.9	14	17	US-10-955-656-7	Sequence 7, Appli	198	42	61.8	9	14	US-10-126-845-21	Sequence 21, Appl
126	53	77.9	14	17	US-10-955-656-65	Sequence 65, Appl	199	42	61.8	9	14	US-10-126-845-79	Sequence 79, Appl
127	53	77.9	20	14	US-10-126-845-110	Sequence 110, App	200	42	61.8	9	15	US-10-116-275-109	Sequence 109, App
128	53	77.9	20	17	US-10-955-656-110	Sequence 110, App	201	42	61.8	9	16	US-10-764-235-21	Sequence 21, Appl
129	51	75.0	17	14	US-10-126-845-24	Sequence 24, Appl	202	42	61.8	9	17	US-10-634-645-7	Sequence 7, Appli
130	51	75.0	17	14	US-10-126-845-82	Sequence 82, Appl	203	42	61.8	9	17	US-10-955-656-21	Sequence 21, Appl
131	51	75.0	17	15	US-10-116-275-112	Sequence 112, App	204	42	61.8	9	17	US-10-955-656-79	Sequence 79, Appl
132	51	75.0	17	16	US-10-764-235-24	Sequence 24, Appl	205	42	61.8	11	14	US-10-126-845-12	Sequence 12, Appl
133	51	75.0	17	17	US-10-955-656-24	Sequence 24, Appl	206	42	61.8	11	14	US-10-126-845-70	Sequence 70, Appl
134	51	75.0	17	17	US-10-955-656-82	Sequence 82, Appl	207	42	61.8	11	15	US-10-116-275-100	Sequence 100, App
135	50	73.5	11	13	US-10-116-288-9	Sequence 9, Appli	208	42	61.8	11	16	US-10-764-235-12	Sequence 12, Appl
136	50	73.5	11	14	US-10-126-845-19	Sequence 19, Appl	209	42	61.8	11	17	US-10-955-656-12	Sequence 12, Appl
137	50	73.5	11	14	US-10-126-845-77	Sequence 77, Appl	210	42	61.8	11	17	US-10-955-656-70	Sequence 70, Appl
138	50	73.5	11	15	US-10-116-275-107	Sequence 107, App	211	42	61.8	102	16	US-10-437-963-110047	Sequence 110047,
139	50	73.5	11	16	US-10-764-235-19	Sequence 19, Appli	212	41	60.3	659	14	US-10-156-761-8777	Sequence 8777, Ap
140	50	73.5	11	17	US-10-634-645-9	Sequence 9, Appli	213	40	58.8	274	15	US-10-424-599-234438	Sequence 234438,
141	50	73.5	11	17	US-10-955-656-19	Sequence 19, Appl	214	40	58.8	364	16	US-10-276-272A-9	Sequence 9, Appli
142	50	73.5	11	17	US-10-955-656-77	Sequence 77, Appl	215	40	58.8	370	15	US-10-425-114-70773	Sequence 70773, A
143	50	73.5	13	14	US-10-126-845-10	Sequence 10, Appl	216	40	58.8	372	16	US-10-276-272A-3	Sequence 3, Appli
144	50	73.5	13	14	US-10-126-845-68	Sequence 68, Appl	217	40	58.8	381	16	US-10-276-272A-10	Sequence 10, Appl
145	50	73.5	13	15	US-10-116-275-98	Sequence 98, Appl	218	40	58.8	389	16	US-10-276-272A-4	Sequence 4, Appli
146	50	73.5	13	16	US-10-764-235-10	Sequence 10, Appl	219	40	58.8	391	16	US-10-276-272A-5	Sequence 5, Appli
147	50	73.5	13	17	US-10-955-656-10	Sequence 10, Appl	220	40	58.8	455	15	US-10-282-122A-54036	Sequence 54036, A
148	50	73.5	13	17	US-10-955-656-68	Sequence 68, Appl	221	40	58.8	458	14	US-10-156-761-12570	Sequence 12570, A
149	47	69.1	11	13	US-10-116-288-5	Sequence 5, Appli	222	40	58.8	582	9	US-09-916-658-4	Sequence 4, Appli
150	47	69.1	11	14	US-10-126-845-15	Sequence 15, Appl	223	40	58.8	582	9	US-09-391-104-28	Sequence 28, Appl
151	47	69.1	11	15	US-10-126-845-73	Sequence 73, Appl	224	40	58.8	582	9	US-09-801-196-27	Sequence 27, Appl
152	47	69.1	11	15	US-10-116-275-103	Sequence 103, App	225	40	58.8	582	9	US-09-919-497-84	Sequence 84, Appl
153	47	69.1	11	16	US-10-764-235-15	Sequence 15, Appl	226	40	58.8	582	10	US-09-916-849A-2	Sequence 2, Appli
154	47	69.1	11	17	US-10-634-645-5	Sequence 5, Appli	227	40	58.8	582	14	US-10-133-797-6	Sequence 6, Appli
155	47	69.1	11	17	US-10-955-656-15	Sequence 15, Appl	228	40	58.8	582	14	US-10-131-985-43	Sequence 43, Appl
156	47	69.1	11	17	US-10-955-656-73	Sequence 73, Appl	229	40	58.8	582	15	US-10-441-010-26	Sequence 26, Appl
157	47	69.1	12	14	US-10-126-845-16	Sequence 16, Appl	230	40	58.8	582	15	US-10-447-315-9	Sequence 9, Appli
158	47	69.1	12	14	US-10-126-845-74	Sequence 74, Appl	231	40	58.8	582	16	US-10-276-272A-19	Sequence 19, Appl
159	47	69.1	12	15	US-10-116-275-104	Sequence 104, App	232	40	58.8	582	17	US-10-901-417-43	Sequence 43, Appl

233	40	58.8	582	17	US-10-953-264-26	Sequence 26, Appl	306	38	55.9	308	17	US-10-732-923-23332	Sequence 23332, A
234	40	58.8	584	16	US-10-276-272A-11	Sequence 11, Appl	307	38	55.9	323	14	US-10-229-567-28	Sequence 28, Appl
235	40	58.8	590	16	US-10-276-272A-2	Sequence 2, Appl	308	38	55.9	337	16	US-10-437-963-113618	Sequence 113618, A
236	40	58.8	592	16	US-10-276-272A-12	Sequence 12, Appl	309	38	55.9	340	14	US-10-259-165-226	Sequence 226, App
237	40	58.8	1246	9	US-09-919-497-85	Sequence 85, Appl	310	38	55.9	340	14	US-10-437-963-171330	Sequence 171330, A
238	40	58.8	1247	10	US-09-910-715-216	Sequence 216, App	311	38	55.9	369	17	US-10-732-923-23329	Sequence 23329, A
239	40	58.8	1247	9	US-09-961-403-14	Sequence 14, Appl	312	38	55.9	370	17	US-10-732-923-23330	Sequence 23330, A
240	39	57.4	9	13	US-10-116-288-3	Sequence 3, Appl	313	38	55.9	370	17	US-10-732-923-23331	Sequence 23331, A
241	39	57.4	9	14	US-10-126-845-17	Sequence 17, Appl	314	38	55.9	377	14	US-10-229-567-29	Sequence 29, Appl
242	39	57.4	9	14	US-10-126-845-75	Sequence 75, Appl	315	38	55.9	377	14	US-10-287-274-450	Sequence 450, App
243	39	57.4	9	15	US-10-116-275-105	Sequence 105, App	316	38	55.9	377	15	US-10-369-493-817	Sequence 817, App
244	39	57.4	9	16	US-10-764-235-17	Sequence 17, Appl	317	38	55.9	377	15	US-10-282-122A-42668	Sequence 42668, A
245	39	57.4	9	17	US-10-634-645-3	Sequence 3, Appl	318	38	55.9	406	16	US-10-767-701-42204	Sequence 42204, A
246	39	57.4	9	17	US-10-955-656-17	Sequence 17, Appl	319	38	55.9	513	14	US-10-156-761-7625	Sequence 7625, Ap
247	39	57.4	9	17	US-10-955-656-75	Sequence 75, Appl	320	38	55.9	532	15	US-10-282-122A-71488	Sequence 71488, A
248	39	57.4	11	14	US-10-126-845-8	Sequence 8, Appl	321	38	55.9	580	16	US-10-408-765A-1733	Sequence 1733, Ap
249	39	57.4	11	14	US-10-126-845-66	Sequence 66, Appl	322	38	55.9	619	14	US-10-168-836-2	Sequence 2, Appl
250	39	57.4	11	15	US-10-116-275-96	Sequence 96, Appl	323	38	55.9	718	16	US-10-437-963-136955	Sequence 136955, A
251	39	57.4	11	16	US-10-764-235-8	Sequence 8, Appl	324	38	55.9	1182	16	US-10-437-963-123209	Sequence 123209, A
252	39	57.4	11	17	US-10-955-656-8	Sequence 8, Appl	325	38	55.9	2802	9	US-09-808-602-81	Sequence 81, Appl
253	39	57.4	11	17	US-10-955-656-66	Sequence 66, Appl	326	38	55.9	2802	10	US-09-800-198-69	Sequence 69, Appl
254	39	57.4	21	14	US-10-126-845-112	Sequence 112, App	327	38	55.9	2802	15	US-10-072-012-489	Sequence 489, App
255	39	57.4	21	17	US-10-955-656-112	Sequence 112, App	328	38	55.9	504	15	US-10-369-493-12595	Sequence 12595, A
256	39	57.4	26	15	US-10-424-599-223184	Sequence 49, Appl	329	38	55.9	504	15	US-10-126-845-85	Sequence 85, Appl
257	39	57.4	93	15	US-10-424-599-223184	Sequence 49, Appl	330	37	54.4	19	14	US-10-126-845-85	Sequence 85, Appl
258	39	57.4	123	16	US-10-437-963-166386	Sequence 166386, A	331	37	54.4	21	15	US-10-955-656-85	Sequence 85, Appl
259	39	57.4	166	16	US-10-437-963-122929	Sequence 122929, A	332	37	54.4	22	15	US-10-449-831A-64	Sequence 64, Appl
260	39	57.4	173	16	US-10-767-701-54937	Sequence 54937, A	333	37	54.4	24	10	US-09-759-1308-327	Sequence 327, App
261	39	57.4	227	14	US-10-235-567A-670	Sequence 670, App	334	37	54.4	24	14	US-10-189-123-57	Sequence 57, Appl
262	39	57.4	235	16	US-10-437-963-202951	Sequence 202951, A	335	37	54.4	24	14	US-10-188-495-57	Sequence 57, Appl
263	39	57.4	240	15	US-10-278-141-5	Sequence 5, Appl	336	37	54.4	24	16	US-10-741-790-327	Sequence 327, App
264	39	57.4	240	15	US-10-296-081-5	Sequence 5, Appl	337	37	54.4	99	11	US-09-833-245-1146	Sequence 1146, Ap
265	39	57.4	254	16	US-10-408-765A-1496	Sequence 1496, Ap	338	37	54.4	117	16	US-10-437-963-203607	Sequence 203607, A
266	39	57.4	273	15	US-10-282-122A-47625	Sequence 47625, A	339	37	54.4	146	11	US-09-833-245-2081	Sequence 2081, Ap
267	39	57.4	273	15	US-10-282-122A-50187	Sequence 50187, A	340	37	54.4	146	14	US-10-060-255-65	Sequence 65, Appl
268	39	57.4	308	9	US-09-894-018-107	Sequence 107, App	341	37	54.4	174	11	US-09-833-245-316	Sequence 316, App
269	39	57.4	383	15	US-10-425-114-49307	Sequence 49307, A	342	37	54.4	198	14	US-10-060-255-79	Sequence 79, Appl
270	39	57.4	425	15	US-10-374-780A-854	Sequence 854, App	343	37	54.4	234	11	US-09-833-245-318	Sequence 318, App
271	39	57.4	425	16	US-10-437-963-189076	Sequence 189076, A	344	37	54.4	307	16	US-10-437-963-104711	Sequence 104711, A
272	39	57.4	481	17	US-10-732-923-1067	Sequence 1067, Ap	345	37	54.4	313	16	US-10-437-963-176697	Sequence 176697, A
273	39	57.4	529	15	US-10-282-122A-70592	Sequence 70592, A	346	37	54.4	321	16	US-10-614-853-28	Sequence 28, Appl
274	39	57.4	534	15	US-10-282-122A-70409	Sequence 70409, A	347	37	54.4	331	15	US-10-371-701-10	Sequence 10, Appl
275	39	57.4	633	15	US-10-333-946-14	Sequence 14, Appl	348	37	54.4	331	15	US-10-371-701-34	Sequence 34, Appl
276	39	57.4	874	9	US-09-963-766-6	Sequence 6, Appl	349	37	54.4	353	9	US-09-828-366-2	Sequence 2, Appl
277	39	57.4	882	15	US-10-104-047-3543	Sequence 3543, Ap	350	37	54.4	353	9	US-09-903-320-2	Sequence 2, Appl
278	39	57.4	1386	9	US-09-866-582-38	Sequence 38, Appl	351	37	54.4	353	9	US-09-903-088B-2	Sequence 2, Appl
279	39	57.4	1386	17	US-10-839-016-38	Sequence 38, Appl	352	37	54.4	353	9	US-09-905-098B-2	Sequence 2, Appl
280	39	57.4	3970	14	US-10-156-761-10429	Sequence 10429, A	353	37	54.4	353	9	US-09-902-853-2	Sequence 2, Appl
281	38	56.6	335	17	US-10-732-923-19017	Sequence 19017, A	354	37	54.4	353	9	US-09-907-824-2	Sequence 2, Appl
282	38	55.9	8	13	US-10-116-288-6	Sequence 6, Appl	355	37	54.4	353	9	US-09-907-841-2	Sequence 2, Appl
283	38	55.9	8	14	US-10-126-845-22	Sequence 22, Appl	356	37	54.4	353	10	US-09-904-011-2	Sequence 2, Appl
284	38	55.9	8	14	US-10-126-845-80	Sequence 80, Appl	357	37	54.4	353	10	US-09-903-840-2	Sequence 2, Appl
285	38	55.9	8	15	US-10-116-275-110	Sequence 110, App	358	37	54.4	353	10	US-09-908-093-2	Sequence 2, Appl
286	38	55.9	8	16	US-10-764-235-22	Sequence 22, Appl	359	37	54.4	353	10	US-09-759-1308-326	Sequence 326, App
287	38	55.9	8	17	US-10-634-645-6	Sequence 6, Appl	360	37	54.4	353	10	US-09-906-742-2	Sequence 2, Appl
288	38	55.9	8	17	US-10-955-656-22	Sequence 22, Appl	361	37	54.4	353	10	US-09-906-838-2	Sequence 2, Appl
289	38	55.9	8	17	US-10-955-656-80	Sequence 80, Appl	362	37	54.4	353	10	US-09-907-613-2	Sequence 2, Appl
290	38	55.9	10	14	US-10-126-845-13	Sequence 13, Appl	363	37	54.4	353	10	US-09-907-942-2	Sequence 2, Appl
291	38	55.9	10	14	US-10-126-845-71	Sequence 71, Appl	364	37	54.4	353	10	US-09-904-859-2	Sequence 2, Appl
292	38	55.9	10	15	US-10-116-275-101	Sequence 101, App	365	37	54.4	353	10	US-09-909-204-2	Sequence 2, Appl
293	38	55.9	10	16	US-10-764-235-13	Sequence 13, Appl	366	37	54.4	353	10	US-09-904-820-2	Sequence 2, Appl
294	38	55.9	10	17	US-10-955-656-13	Sequence 13, Appl	367	37	54.4	353	10	US-09-904-786-2	Sequence 2, Appl
295	38	55.9	10	17	US-10-955-656-71	Sequence 71, Appl	368	37	54.4	353	10	US-09-906-646-2	Sequence 2, Appl
296	38	55.9	130	16	US-10-767-701-59556	Sequence 59556, A	369	37	54.4	353	10	US-09-906-700-2	Sequence 2, Appl
297	38	55.9	140	15	US-10-424-599-232886	Sequence 232886, A	370	37	54.4	353	10	US-09-903-786-2	Sequence 2, Appl
298	38	55.9	153	16	US-10-767-701-51960	Sequence 51960, A	371	37	54.4	353	10	US-09-902-903-2	Sequence 2, Appl
299	38	55.9	159	16	US-10-767-701-36783	Sequence 36783, A	372	37	54.4	353	10	US-09-903-749A-2	Sequence 2, Appl
300	38	55.9	162	10	US-09-826-734-112	Sequence 112, App	373	37	54.4	353	10	US-09-904-119-2	Sequence 2, Appl
301	38	55.9	162	10	US-09-826-734-118	Sequence 118, App	374	37	54.4	353	10	US-09-904-956-2	Sequence 2, Appl
302	38	55.9	182	15	US-10-369-493-22057	Sequence 22057, A	375	37	54.4	353	10	US-09-902-736-2	Sequence 2, Appl
303	38	55.9	248	15	US-10-115-479-18	Sequence 18, Appl	376	37	54.4	353	10	US-09-907-794-2	Sequence 2, Appl
304	38	55.9	295	15	US-10-425-114-60696	Sequence 60696, A	377	37	54.4	353	10	US-09-903-943-2	Sequence 2, Appl
305	38	55.9	305	15	US-10-282-122A-63318	Sequence 63318, A	378	37	54.4	353	10	US-09-904-462-2	Sequence 2, Appl

379	37	54.4	353	10	US-09-907-925-2	Sequence 2, Appli	452	37	54.4	353	14	US-10-127-825A-296	Sequence 296, App
380	37	54.4	353	10	US-09-902-692-2	Sequence 2, Appli	453	37	54.4	353	14	US-10-127-829A-296	Sequence 296, App
381	37	54.4	353	10	US-09-903-520-2	Sequence 2, Appli	454	37	54.4	353	14	US-10-127-835A-296	Sequence 296, App
382	37	54.4	353	10	US-09-905-056-2	Sequence 2, Appli	455	37	54.4	353	14	US-10-127-839A-296	Sequence 296, App
383	37	54.4	353	10	US-09-909-064-2	Sequence 2, Appli	456	37	54.4	353	14	US-10-127-901A-296	Sequence 296, App
384	37	54.4	353	10	US-09-904-553-2	Sequence 2, Appli	457	37	54.4	353	14	US-10-128-693A-296	Sequence 296, App
385	37	54.4	353	10	US-09-905-381-2	Sequence 2, Appli	458	37	54.4	353	14	US-10-131-813A-296	Sequence 296, App
386	37	54.4	353	10	US-09-904-485-2	Sequence 2, Appli	459	37	54.4	353	14	US-10-131-818A-296	Sequence 296, App
387	37	54.4	353	10	US-09-905-348-2	Sequence 2, Appli	460	37	54.4	353	14	US-10-131-823A-296	Sequence 296, App
388	37	54.4	353	10	US-09-905-088-2	Sequence 2, Appli	461	37	54.4	353	14	US-10-131-824A-296	Sequence 296, App
389	37	54.4	353	10	US-09-907-575-2	Sequence 2, Appli	462	37	54.4	353	14	US-10-131-830A-296	Sequence 296, App
390	37	54.4	353	10	US-09-905-075-2	Sequence 2, Appli	463	37	54.4	353	14	US-10-131-837A-296	Sequence 296, App
391	37	54.4	353	10	US-09-902-759-2	Sequence 2, Appli	464	37	54.4	353	14	US-10-137-872A-296	Sequence 296, App
392	37	54.4	353	10	US-09-902-634-2	Sequence 2, Appli	465	37	54.4	353	14	US-10-147-500-296	Sequence 296, App
393	37	54.4	353	10	US-09-902-713-2	Sequence 2, Appli	466	37	54.4	353	14	US-10-147-502-296	Sequence 296, App
394	37	54.4	353	10	US-09-907-979-2	Sequence 2, Appli	467	37	54.4	353	14	US-10-147-515-296	Sequence 296, App
395	37	54.4	353	10	US-09-902-615-2	Sequence 2, Appli	468	37	54.4	353	14	US-10-147-517-296	Sequence 296, App
396	37	54.4	353	10	US-09-903-925-2	Sequence 2, Appli	469	37	54.4	353	14	US-10-147-526-296	Sequence 296, App
397	37	54.4	353	10	US-09-906-760A-2	Sequence 2, Appli	470	37	54.4	353	14	US-10-147-527-296	Sequence 296, App
398	37	54.4	353	10	US-09-903-823-2	Sequence 2, Appli	471	37	54.4	353	14	US-10-121-041-296	Sequence 296, App
399	37	54.4	353	10	US-09-907-652-2	Sequence 2, Appli	472	37	54.4	353	14	US-10-121-043-296	Sequence 296, App
400	37	54.4	353	10	US-09-902-572A-2	Sequence 2, Appli	473	37	54.4	353	14	US-10-121-047-296	Sequence 296, App
401	37	54.4	353	10	US-09-902-979-2	Sequence 2, Appli	474	37	54.4	353	14	US-10-123-215-296	Sequence 296, App
402	37	54.4	353	10	US-09-905-125-2	Sequence 2, Appli	475	37	54.4	353	14	US-10-123-902-296	Sequence 296, App
403	37	54.4	353	10	US-09-906-815A-2	Sequence 2, Appli	476	37	54.4	353	14	US-10-123-908-296	Sequence 296, App
404	37	54.4	353	10	US-09-905-449-2	Sequence 2, Appli	477	37	54.4	353	14	US-10-123-909-296	Sequence 296, App
405	37	54.4	353	10	US-09-903-806-2	Sequence 2, Appli	478	37	54.4	353	14	US-10-123-910-296	Sequence 296, App
406	37	54.4	353	10	US-09-904-992-2	Sequence 2, Appli	479	37	54.4	353	14	US-10-124-813-296	Sequence 296, App
407	37	54.4	353	10	US-09-904-838-2	Sequence 2, Appli	480	37	54.4	353	14	US-10-124-817-296	Sequence 296, App
408	37	54.4	353	10	US-09-906-777-2	Sequence 2, Appli	481	37	54.4	353	14	US-10-125-922-296	Sequence 296, App
409	37	54.4	353	10	US-09-903-603A-2	Sequence 2, Appli	482	37	54.4	353	14	US-10-125-924-296	Sequence 296, App
410	37	54.4	353	10	US-09-904-532-2	Sequence 2, Appli	483	37	54.4	353	14	US-10-140-860-296	Sequence 296, App
411	37	54.4	353	10	US-09-904-766-2	Sequence 2, Appli	484	37	54.4	353	14	US-10-142-417-296	Sequence 296, App
412	37	54.4	353	10	US-09-904-920A-2	Sequence 2, Appli	485	37	54.4	353	14	US-10-147-519-296	Sequence 296, App
413	37	54.4	353	10	US-09-904-877A-2	Sequence 2, Appli	486	37	54.4	353	14	US-10-157-782-296	Sequence 296, App
414	37	54.4	353	10	US-09-903-562-2	Sequence 2, Appli	487	37	54.4	353	14	US-10-152-395-296	Sequence 296, App
415	37	54.4	353	10	US-09-906-618-2	Sequence 2, Appli	488	37	54.4	353	14	US-10-189-123-56	Sequence 56, Appl
416	37	54.4	353	10	US-09-907-728-2	Sequence 2, Appli	489	37	54.4	353	14	US-10-125-926A-296	Sequence 296, App
417	37	54.4	353	10	US-09-904-805-2	Sequence 2, Appli	490	37	54.4	353	14	US-10-125-930A-296	Sequence 296, App
418	37	54.4	353	10	US-09-904-938A-2	Sequence 2, Appli	491	37	54.4	353	14	US-10-127-831A-296	Sequence 296, App
419	37	54.4	353	10	US-09-906-722A-2	Sequence 2, Appli	492	37	54.4	353	14	US-10-127-837A-296	Sequence 296, App
420	37	54.4	353	11	US-09-908-576-2	Sequence 2, Appli	493	37	54.4	353	14	US-10-127-838B-296	Sequence 296, App
421	37	54.4	353	14	US-10-028-072-296	Sequence 296, App	494	37	54.4	353	14	US-10-127-842A-296	Sequence 296, App
422	37	54.4	353	14	US-10-140-808-296	Sequence 296, App	495	37	54.4	353	14	US-10-127-843A-296	Sequence 296, App
423	37	54.4	353	14	US-10-121-049-296	Sequence 296, App	496	37	54.4	353	14	US-10-127-845A-296	Sequence 296, App
424	37	54.4	353	14	US-10-123-904-296	Sequence 296, App	497	37	54.4	353	14	US-10-127-846A-296	Sequence 296, App
425	37	54.4	353	14	US-10-140-470-296	Sequence 296, App	498	37	54.4	353	14	US-10-127-848A-296	Sequence 296, App
426	37	54.4	353	14	US-10-175-746-296	Sequence 296, App	499	37	54.4	353	14	US-10-127-849A-296	Sequence 296, App
427	37	54.4	353	14	US-10-176-918-296	Sequence 296, App	500	37	54.4	353	14	US-10-127-850A-296	Sequence 296, App
428	37	54.4	353	14	US-10-176-921-296	Sequence 296, App	501	37	54.4	353	14	US-10-127-851A-296	Sequence 296, App
429	37	54.4	353	14	US-10-137-865-296	Sequence 296, App	502	37	54.4	353	14	US-10-128-684A-296	Sequence 296, App
430	37	54.4	353	14	US-10-140-474-296	Sequence 296, App	503	37	54.4	353	14	US-10-128-686A-296	Sequence 296, App
431	37	54.4	353	14	US-10-142-431-296	Sequence 296, App	504	37	54.4	353	14	US-10-128-690A-296	Sequence 296, App
432	37	54.4	353	14	US-10-143-114-296	Sequence 296, App	505	37	54.4	353	14	US-10-128-691A-296	Sequence 296, App
433	37	54.4	353	14	US-10-142-419-296	Sequence 296, App	506	37	54.4	353	14	US-10-131-819A-296	Sequence 296, App
434	37	54.4	353	14	US-10-123-262-296	Sequence 296, App	507	37	54.4	353	14	US-10-131-829A-296	Sequence 296, App
435	37	54.4	353	14	US-10-142-423-296	Sequence 296, App	508	37	54.4	353	14	US-10-131-836A-296	Sequence 296, App
436	37	54.4	353	14	US-10-121-050-296	Sequence 296, App	509	37	54.4	353	14	US-10-146-729-296	Sequence 296, App
437	37	54.4	353	14	US-10-141-755-296	Sequence 296, App	510	37	54.4	353	14	US-10-146-791-296	Sequence 296, App
438	37	54.4	353	14	US-10-143-032-296	Sequence 296, App	511	37	54.4	353	14	US-10-147-484-296	Sequence 296, App
439	37	54.4	353	14	US-10-123-108-296	Sequence 296, App	512	37	54.4	353	14	US-10-147-508-296	Sequence 296, App
440	37	54.4	353	14	US-10-123-236-296	Sequence 296, App	513	37	54.4	353	14	US-10-147-512-296	Sequence 296, App
441	37	54.4	353	14	US-10-123-261-296	Sequence 296, App	514	37	54.4	353	14	US-10-175-735-296	Sequence 296, App
442	37	54.4	353	14	US-10-123-921-296	Sequence 296, App	515	37	54.4	353	14	US-10-121-040-296	Sequence 296, App
443	37	54.4	353	14	US-10-140-928-296	Sequence 296, App	516	37	54.4	353	14	US-10-121-056-296	Sequence 296, App
444	37	54.4	353	14	US-10-121-045-296	Sequence 296, App	517	37	54.4	353	14	US-10-121-061-296	Sequence 296, App
445	37	54.4	353	14	US-10-123-292-296	Sequence 296, App	518	37	54.4	353	14	US-10-123-235-296	Sequence 296, App
446	37	54.4	353	14	US-10-123-903-296	Sequence 296, App	519	37	54.4	353	14	US-10-124-818-296	Sequence 296, App
447	37	54.4	353	14	US-10-124-819-296	Sequence 296, App	520	37	54.4	353	14	US-10-137-868-296	Sequence 296, App
448	37	54.4	353	14	US-10-124-822-296	Sequence 296, App	521	37	54.4	353	14	US-10-147-452-296	Sequence 296, App
449	37	54.4	353	14	US-10-140-925-296	Sequence 296, App	522	37	54.4	353	14	US-10-158-782-296	Sequence 296, App
450	37	54.4	353	14	US-10-160-498-296	Sequence 296, App	523	37	54.4	353	14	US-10-123-905-296	Sequence 296, App
451	37	54.4	353	14	US-10-124-824-296	Sequence 296, App	524	37	54.4	353	14	US-10-123-907-296	Sequence 296, App



```
671 37 54.4 353 14 US-10-147-480-296 Sequence 296, App
672 37 54.4 353 14 US-10-147-485-296 Sequence 296, App
673 37 54.4 353 14 US-10-147-486-296 Sequence 296, App
674 37 54.4 353 14 US-10-147-487-296 Sequence 296, App
675 37 54.4 353 14 US-10-147-490-296 Sequence 296, App
676 37 54.4 353 14 US-10-147-494-296 Sequence 296, App
677 37 54.4 353 14 US-10-147-498-296 Sequence 296, App
678 37 54.4 353 14 US-10-147-514-296 Sequence 296, App
679 37 54.4 353 14 US-10-147-524-296 Sequence 296, App
680 37 54.4 353 14 US-10-152-379-296 Sequence 296, App
681 37 54.4 353 14 US-10-152-394-296 Sequence 296, App
682 37 54.4 353 14 US-10-152-406-296 Sequence 296, App
683 37 54.4 353 14 US-10-156-847-296 Sequence 296, App
684 37 54.4 353 14 US-10-157-778-296 Sequence 296, App
685 37 54.4 353 14 US-10-157-789-296 Sequence 296, App
686 37 54.4 353 14 US-10-160-504-296 Sequence 296, App
687 37 54.4 353 14 US-10-145-634-296 Sequence 296, App
688 37 54.4 353 14 US-10-147-520-296 Sequence 296, App
689 37 54.4 353 14 US-10-157-781-296 Sequence 296, App
690 37 54.4 353 14 US-10-176-989-296 Sequence 296, App
691 37 54.4 353 14 US-10-188-485-56 Sequence 56, Appl
692 37 54.4 353 14 US-10-147-491-296 Sequence 296, App
693 37 54.4 353 14 US-10-152-378-296 Sequence 296, App
694 37 54.4 353 14 US-10-152-382-296 Sequence 296, App
695 37 54.4 353 14 US-10-152-383-296 Sequence 296, App
696 37 54.4 353 14 US-10-152-384-296 Sequence 296, App
697 37 54.4 353 14 US-10-152-387-296 Sequence 296, App
698 37 54.4 353 14 US-10-152-389-296 Sequence 296, App
699 37 54.4 353 14 US-10-152-390-296 Sequence 296, App
700 37 54.4 353 14 US-10-152-392-296 Sequence 296, App
701 37 54.4 353 14 US-10-153-756-296 Sequence 296, App
702 37 54.4 353 14 US-10-157-784-296 Sequence 296, App
703 37 54.4 353 14 US-10-157-797-296 Sequence 296, App
704 37 54.4 353 14 US-10-158-491-296 Sequence 296, App
705 37 54.4 353 14 US-10-299-976-2 Sequence 2, Appli
706 37 54.4 353 14 US-10-142-762-296 Sequence 296, App
707 37 54.4 353 14 US-10-142-764-296 Sequence 296, App
708 37 54.4 353 14 US-10-142-766-296 Sequence 296, App
709 37 54.4 353 14 US-10-145-625-296 Sequence 296, App
710 37 54.4 353 14 US-10-145-627-296 Sequence 296, App
711 37 54.4 353 14 US-10-145-960-296 Sequence 296, App
712 37 54.4 353 14 US-10-145-962-296 Sequence 296, App
713 37 54.4 353 14 US-10-146-789-296 Sequence 296, App
714 37 54.4 353 14 US-10-147-483-296 Sequence 296, App
715 37 54.4 353 14 US-10-147-496-296 Sequence 296, App
716 37 54.4 353 14 US-10-147-505-296 Sequence 296, App
717 37 54.4 353 14 US-10-147-516-296 Sequence 296, App
718 37 54.4 353 14 US-10-152-398-296 Sequence 296, App
719 37 54.4 353 14 US-10-139-980-296 Sequence 2, Appli
720 37 54.4 353 14 US-10-299-937-2 Sequence 296, App
721 37 54.4 353 14 US-10-145-750-296 Sequence 296, App
722 37 54.4 353 14 US-10-152-373-296 Sequence 296, App
723 37 54.4 353 14 US-10-121-044-296 Sequence 296, App
724 37 54.4 353 14 US-10-121-055-296 Sequence 296, App
725 37 54.4 353 14 US-10-121-057-296 Sequence 296, App
726 37 54.4 353 14 US-10-121-058-296 Sequence 296, App
727 37 54.4 353 14 US-10-121-059-296 Sequence 296, App
728 37 54.4 353 14 US-10-121-060-296 Sequence 296, App
729 37 54.4 353 14 US-10-123-109-296 Sequence 296, App
730 37 54.4 353 14 US-10-123-154-296 Sequence 296, App
731 37 54.4 353 14 US-10-123-157-296 Sequence 296, App
732 37 54.4 353 14 US-10-123-906-296 Sequence 296, App
733 37 54.4 353 14 US-10-124-814-296 Sequence 296, App
734 37 54.4 353 14 US-10-124-816-296 Sequence 296, App
735 37 54.4 353 14 US-10-124-820-296 Sequence 296, App
736 37 54.4 353 14 US-10-125-704-296 Sequence 296, App
737 37 54.4 353 14 US-10-125-927-296 Sequence 296, App
738 37 54.4 353 14 US-10-142-889-296 Sequence 296, App
739 37 54.4 353 14 US-10-145-874-296 Sequence 296, App
740 37 54.4 353 14 US-10-147-497-296 Sequence 296, App
741 37 54.4 353 14 US-10-152-371-296 Sequence 296, App
742 37 54.4 353 14 US-10-152-374-296 Sequence 296, App
743 37 54.4 353 14 US-10-152-375-296 Sequence 296, App
744 37 54.4 353 14 US-10-152-377-296 Sequence 296, App
745 37 54.4 353 14 US-10-152-386-296 Sequence 296, App
746 37 54.4 353 14 US-10-152-391-296 Sequence 296, App
747 37 54.4 353 14 US-10-152-399-296 Sequence 296, App
748 37 54.4 353 14 US-10-156-848-296 Sequence 296, App
749 37 54.4 353 14 US-10-157-785-296 Sequence 296, App
750 37 54.4 353 14 US-10-157-794-296 Sequence 296, App

ALIGNMENTS

RESULT 1
US-10-126-845-2
; Sequence 2, Application US/10126845
; Publication No. US20030181367A1
; GENERAL INFORMATION:
; APPLICANT: O'Mahony, Daniel J.
; APPLICANT: Lambkin, Imelda J.
; APPLICANT: Pinilla, Clemencia
; APPLICANT: Houghten, Richard
; TITLE OF INVENTION: MEMBRANE TRANSLOCATING PEPTIDE DRUG DELIVERY SYSTEM
; FILE REFERENCE: E1067/20058
; CURRENT APPLICATION NUMBER: US/10/126,845
; CURRENT FILING DATE: 2002-10-15
; NUMBER OF SEQ ID NOS: 119
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: membrane translocating peptide
US-10-126-845-2
Query Match 100.0%; Score 68; DB 14; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.00029;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKAADVLLPVLAAAP 15
|||||
DB 1 KKAADVLLPVLAAAP 15

RESULT 2
US-10-126-845-3
; Sequence 3, Application US/10126845
; Publication No. US20030181367A1
; GENERAL INFORMATION:
; APPLICANT: O'Mahony, Daniel J.
; APPLICANT: Lambkin, Imelda J.
; APPLICANT: Pinilla, Clemencia
; APPLICANT: Houghten, Richard
; TITLE OF INVENTION: MEMBRANE TRANSLOCATING PEPTIDE DRUG DELIVERY SYSTEM
; FILE REFERENCE: E1067/20058
; CURRENT APPLICATION NUMBER: US/10/126,845
; CURRENT FILING DATE: 2002-10-15
; NUMBER OF SEQ ID NOS: 119
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: membrane translocating peptide
; NAME/KEY: MOD RES
; LOCATION: (15)..(15)
; OTHER INFORMATION: linked to FITC-LC
US-10-126-845-3
Query Match 100.0%; Score 68; DB 14; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.00029;
```



Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KKAARVLLPVLLAAP 15  
| | | | | | | | | | | | | | |  
Db 1 KKAARVLLPVLLAAP 15

RESULT 3  
US-10-126-845-60  
; Sequence 60, Application US/10126845  
; Publication No. US20030181367A1  
; GENERAL INFORMATION:  
; APPLICANT: O'Mahony, Daniel J.  
; APPLICANT: Lambkin, Imelda J.  
; APPLICANT: Pillila, Clemencia  
; APPLICANT: Houghten, Richard  
; TITLE OF INVENTION: MEMBRANE TRANSLOCATING PEPTIDE DRUG DELIVERY SYSTEM  
; FILE REFERENCE: E1067/20058  
; CURRENT APPLICATION NUMBER: US/10/126,845  
; CURRENT FILING DATE: 2002-10-15  
; NUMBER OF SEQ ID NOS: 119  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 60  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; FEATURE:  
; OTHER INFORMATION: D form peptide  
; NAME/KEY: MISC\_FEATURE  
; LOCATION: (1)..(15)  
; OTHER INFORMATION: D form amino acid  
US-10-126-845-60

Query Match 100.0%; Score 68; DB 14; Length 15;  
Best Local Similarity 100.0%; Pred. No. 0.00029;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KKAARVLLPVLLAAP 15  
| | | | | | | | | | | | | | |  
Db 1 KKAARVLLPVLLAAP 15

RESULT 4  
US-10-126-845-61  
; Sequence 61, Application US/10126845  
; Publication No. US20030181367A1  
; GENERAL INFORMATION:  
; APPLICANT: O'Mahony, Daniel J.  
; APPLICANT: Lambkin, Imelda J.  
; APPLICANT: Pillila, Clemencia  
; APPLICANT: Houghten, Richard  
; TITLE OF INVENTION: MEMBRANE TRANSLOCATING PEPTIDE DRUG DELIVERY SYSTEM  
; FILE REFERENCE: E1067/20058  
; CURRENT APPLICATION NUMBER: US/10/126,845  
; CURRENT FILING DATE: 2002-10-15  
; NUMBER OF SEQ ID NOS: 119  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 61  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: D form peptide  
; NAME/KEY: MISC\_FEATURE  
; LOCATION: (1)..(15)  
; OTHER INFORMATION: D form amino acid  
US-10-126-845-61

Query Match 100.0%; Score 68; DB 14; Length 15;  
Best Local Similarity 100.0%; Pred. No. 0.00029;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KKAARVLLPVLLAAP 15  
| | | | | | | | | | | | | | |  
Db 1 KKAARVLLPVLLAAP 15

RESULT 5  
US-10-136-187-2  
; Sequence 2, Application US/10136187  
; Publication No. US20030203865A1  
; GENERAL INFORMATION:  
; APPLICANT: Harvie, Pierrot  
; APPLICANT: Paul, Ralph  
; APPLICANT: Cudmore, Sally  
; APPLICANT: O'Mahony, Daniel J.  
; TITLE OF INVENTION: LIPID-COMPRISING DRUG DELIVERY COMPLEXES  
; TITLE OF INVENTION: AND METHODS FOR THEIR PRODUCTION  
; FILE REFERENCE: 226272005300  
; CURRENT APPLICATION NUMBER: US/10/136,187  
; CURRENT FILING DATE: 2002-09-13  
; PRIOR APPLICATION NUMBER: US 60/287,786  
; PRIOR FILING DATE: 2001-04-30  
; NUMBER OF SEQ ID NOS: 45  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Synthetic Construct  
US-10-136-187-2

Query Match 100.0%; Score 68; DB 15; Length 15;  
Best Local Similarity 100.0%; Pred. No. 0.00029;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KKAARVLLPVLLAAP 15  
| | | | | | | | | | | | | | |  
Db 1 KKAARVLLPVLLAAP 15

RESULT 6  
US-10-136-187-6  
; Sequence 6, Application US/10136187  
; Publication No. US20030203865A1  
; GENERAL INFORMATION:  
; APPLICANT: Harvie, Pierrot  
; APPLICANT: Paul, Ralph  
; APPLICANT: Cudmore, Sally  
; APPLICANT: O'Mahony, Daniel J.  
; TITLE OF INVENTION: LIPID-COMPRISING DRUG DELIVERY COMPLEXES  
; TITLE OF INVENTION: AND METHODS FOR THEIR PRODUCTION  
; FILE REFERENCE: 226272005300  
; CURRENT APPLICATION NUMBER: US/10/136,187  
; CURRENT FILING DATE: 2002-09-13  
; PRIOR APPLICATION NUMBER: US 60/287,786  
; PRIOR FILING DATE: 2001-04-30  
; NUMBER OF SEQ ID NOS: 45  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 6  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Synthetic Construct  
; NAME/KEY: VARIANT  
; LOCATION: 1  
; OTHER INFORMATION: Lysine is attached to galactose  
US-10-136-187-6

Query Match 100.0%; Score 68; DB 15; Length 15;  
Best Local Similarity 100.0%; Pred. No. 0.00029;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKA...VLLPVLLAAP 15  
 Db 1 KKA...VLLPVLLAAP 15

RESULT 7

US-10-136-187-12  
 ; Sequence 12, Application US/10136187  
 ; Publication No. US20030203865A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Harvie, Pierrot  
 ; APPLICANT: ~~Paul~~, Ralph  
 ; APPLICANT: ~~Cudmore~~, Sally  
 ; APPLICANT: O'Mahony, Daniel J.  
 ; TITLE OF INVENTION: LIPID-COMPRISING DRUG DELIVERY COMPLEXES  
 ; TITLE OF INVENTION: AND METHODS FOR THEIR PRODUCTION  
 ; FILE REFERENCE: 226272005300  
 ; CURRENT APPLICATION NUMBER: US/10/136,187  
 ; CURRENT FILING DATE: 2002-09-13  
 ; PRIOR APPLICATION NUMBER: US 60/287,786  
 ; PRIOR FILING DATE: 2001-04-30  
 ; NUMBER OF SEQ ID NOS: 45  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 12  
 ; LENGTH: 15  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Synthetic Construct  
 ; NAME/KEY: VARIANT  
 ; LOCATION: 1  
 ; OTHER INFORMATION: Lysine is attached to DOPE-succinyl  
 US-10-136-187-12

Query Match 100.0%; Score 68; DB 15; Length 15;  
 Best Local Similarity 100.0%; Pred. No. 0.00029;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKA...VLLPVLLAAP 15  
 Db 1 KKA...VLLPVLLAAP 15

RESULT 8

US-10-136-187-13  
 ; Sequence 13, Application US/10136187  
 ; Publication No. US20030203865A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Harvie, Pierrot  
 ; APPLICANT: ~~Paul~~, Ralph  
 ; APPLICANT: ~~Cudmore~~, Sally  
 ; APPLICANT: O'Mahony, Daniel J.  
 ; TITLE OF INVENTION: LIPID-COMPRISING DRUG DELIVERY COMPLEXES  
 ; TITLE OF INVENTION: AND METHODS FOR THEIR PRODUCTION  
 ; FILE REFERENCE: 226272005300  
 ; CURRENT APPLICATION NUMBER: US/10/136,187  
 ; CURRENT FILING DATE: 2002-09-13  
 ; PRIOR APPLICATION NUMBER: US 60/287,786  
 ; PRIOR FILING DATE: 2001-04-30  
 ; NUMBER OF SEQ ID NOS: 45  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 13  
 ; LENGTH: 15  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Synthetic Construct  
 US-10-136-187-13

Query Match 100.0%; Score 68; DB 15; Length 15;

Best Local Similarity 100.0%; Pred. No. 0.00029;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKA...VLLPVLLAAP 15  
 Db 1 KKA...VLLPVLLAAP 15

RESULT 9

US-10-136-187-33  
 ; Sequence 33, Application US/10136187  
 ; Publication No. US20030203865A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Harvie, Pierrot  
 ; APPLICANT: ~~Paul~~, Ralph  
 ; APPLICANT: ~~Cudmore~~, Sally  
 ; APPLICANT: O'Mahony, Daniel J.  
 ; TITLE OF INVENTION: LIPID-COMPRISING DRUG DELIVERY COMPLEXES  
 ; TITLE OF INVENTION: AND METHODS FOR THEIR PRODUCTION  
 ; FILE REFERENCE: 226272005300  
 ; CURRENT APPLICATION NUMBER: US/10/136,187  
 ; CURRENT FILING DATE: 2002-09-13  
 ; PRIOR APPLICATION NUMBER: US 60/287,786  
 ; PRIOR FILING DATE: 2001-04-30  
 ; NUMBER OF SEQ ID NOS: 45  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 33  
 ; LENGTH: 15  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Synthetic Construct  
 ; NAME/KEY: VARIANT  
 ; LOCATION: 1  
 ; OTHER INFORMATION: Lysine is attached to Cholesteryl-succinyl  
 US-10-136-187-33

Query Match 100.0%; Score 68; DB 15; Length 15;  
 Best Local Similarity 100.0%; Pred. No. 0.00029;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKA...VLLPVLLAAP 15  
 Db 1 KKA...VLLPVLLAAP 15

RESULT 10

US-10-136-187-34  
 ; Sequence 34, Application US/10136187  
 ; Publication No. US20030203865A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Harvie, Pierrot  
 ; APPLICANT: ~~Paul~~, Ralph  
 ; APPLICANT: ~~Cudmore~~, Sally  
 ; APPLICANT: O'Mahony, Daniel J.  
 ; TITLE OF INVENTION: LIPID-COMPRISING DRUG DELIVERY COMPLEXES  
 ; TITLE OF INVENTION: AND METHODS FOR THEIR PRODUCTION  
 ; FILE REFERENCE: 226272005300  
 ; CURRENT APPLICATION NUMBER: US/10/136,187  
 ; CURRENT FILING DATE: 2002-09-13  
 ; PRIOR APPLICATION NUMBER: US 60/287,786  
 ; PRIOR FILING DATE: 2001-04-30  
 ; NUMBER OF SEQ ID NOS: 45  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 34  
 ; LENGTH: 15  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Synthetic Construct  
 ; NAME/KEY: VARIANT

; LOCATION: 1  
; OTHER INFORMATION: Lysine is attached to DOPE-succinyl  
US-10-136-187-34

Query Match 100.0%; Score 68; DB 15; Length 15;  
Best Local Similarity 100.0%; Pred. No. 0.00029;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KKAARVLLPVLLAAP 15  
| | | | | | | | | | | | | | |  
Db 1 KKAARVLLPVLLAAP 15

RESULT 11  
US-10-136-187-35  
; Sequence 35, Application US/10136187  
; Publication No. US20030203865A1  
; GENERAL INFORMATION:  
; APPLICANT: Harvie, Pierrot  
; APPLICANT: Paul, Ralph  
; APPLICANT: O'Mahony, Sally  
; APPLICANT: O'Mahony, Daniel J.  
; TITLE OF INVENTION: LIPID-COMPRISING DRUG DELIVERY COMPLEXES  
; TITLE OF INVENTION: AND METHODS FOR THEIR PRODUCTION  
; FILE REFERENCE: 226272005300  
; CURRENT APPLICATION NUMBER: US/10/136,187  
; CURRENT FILING DATE: 2002-09-13  
; PRIOR APPLICATION NUMBER: US 60/287,786  
; PRIOR FILING DATE: 2001-04-30  
; NUMBER OF SEQ ID NOS: 45  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 35  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Synthetic Construct  
; NAME/KEY: VARIANT  
; LOCATION: 1  
; OTHER INFORMATION: Lysine is attached to Cholesteryl-succinyl  
US-10-136-187-35

Query Match 100.0%; Score 68; DB 15; Length 15;  
Best Local Similarity 100.0%; Pred. No. 0.00029;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KKAARVLLPVLLAAP 15  
| | | | | | | | | | | | | | |  
Db 1 KKAARVLLPVLLAAP 15

RESULT 12  
US-10-136-187-36  
; Sequence 36, Application US/10136187  
; Publication No. US20030203865A1  
; GENERAL INFORMATION:  
; APPLICANT: Harvie, Pierrot  
; APPLICANT: Paul, Ralph  
; APPLICANT: O'Mahony, Sally  
; APPLICANT: O'Mahony, Daniel J.  
; TITLE OF INVENTION: LIPID-COMPRISING DRUG DELIVERY COMPLEXES  
; TITLE OF INVENTION: AND METHODS FOR THEIR PRODUCTION  
; FILE REFERENCE: 226272005300  
; CURRENT APPLICATION NUMBER: US/10/136,187  
; CURRENT FILING DATE: 2002-09-13  
; PRIOR APPLICATION NUMBER: US 60/287,786  
; PRIOR FILING DATE: 2001-04-30  
; NUMBER OF SEQ ID NOS: 45

; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 36  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Synthetic Construct  
; FEATURE:  
; NAME/KEY: VARIANT  
; LOCATION: 1  
; OTHER INFORMATION: Lysine is attached to DSPE-PEG5K-succinyl  
US-10-136-187-36

Query Match 100.0%; Score 68; DB 15; Length 15;  
Best Local Similarity 100.0%; Pred. No. 0.00029;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KKAARVLLPVLLAAP 15  
| | | | | | | | | | | | | | |  
Db 1 KKAARVLLPVLLAAP 15

RESULT 13  
US-10-136-187-37  
; Sequence 37, Application US/10136187  
; Publication No. US20030203865A1  
; GENERAL INFORMATION:  
; APPLICANT: Harvie, Pierrot  
; APPLICANT: Paul, Ralph  
; APPLICANT: O'Mahony, Sally  
; APPLICANT: O'Mahony, Daniel J.  
; TITLE OF INVENTION: LIPID-COMPRISING DRUG DELIVERY COMPLEXES  
; TITLE OF INVENTION: AND METHODS FOR THEIR PRODUCTION  
; FILE REFERENCE: 226272005300  
; CURRENT APPLICATION NUMBER: US/10/136,187  
; CURRENT FILING DATE: 2002-09-13  
; PRIOR APPLICATION NUMBER: US 60/287,786  
; PRIOR FILING DATE: 2001-04-30  
; NUMBER OF SEQ ID NOS: 45  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 37  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Synthetic Construct  
; NAME/KEY: VARIANT  
; LOCATION: 1  
; OTHER INFORMATION: Lysine is attached to DMPE-PEG5K-succinyl  
US-10-136-187-37

Query Match 100.0%; Score 68; DB 15; Length 15;  
Best Local Similarity 100.0%; Pred. No. 0.00029;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KKAARVLLPVLLAAP 15  
| | | | | | | | | | | | | | |  
Db 1 KKAARVLLPVLLAAP 15

RESULT 14  
US-10-136-187-38  
; Sequence 38, Application US/10136187  
; Publication No. US20030203865A1  
; GENERAL INFORMATION:  
; APPLICANT: Harvie, Pierrot  
; APPLICANT: Paul, Ralph  
; APPLICANT: O'Mahony, Sally  
; APPLICANT: O'Mahony, Daniel J.  
; TITLE OF INVENTION: LIPID-COMPRISING DRUG DELIVERY COMPLEXES  
; TITLE OF INVENTION: AND METHODS FOR THEIR PRODUCTION  
; FILE REFERENCE: 226272005300

; CURRENT APPLICATION NUMBER: US/10/136,187  
; CURRENT FILING DATE: 2002-09-13  
; PRIOR APPLICATION NUMBER: US 60/287,786  
; PRIOR FILING DATE: 2001-04-30  
; NUMBER OF SEQ ID NOS: 45  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 38  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Synthetic Construct  
; NAME/KEY: VARIANT  
; LOCATION: 1  
; OTHER INFORMATION: Lysine is attached to DSPE-PEG5K-succinyl  
US-10-136-187-38

Query Match 100.0%; Score 68; DB 15; Length 15;  
Best Local Similarity 100.0%; Pred. No. 0.00029;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKAADVLLPVLLAAP 15  
|||||  
Db 1 KKAADVLLPVLLAAP 15

## RESULT 15

US-10-136-187-39  
; Sequence 39, Application US/10136187  
; Publication No. US20030203865A1  
; GENERAL INFORMATION:  
; APPLICANT: Harvie, Pierrot  
; APPLICANT: Paul, Ralph  
; APPLICANT: Cudmore, Sally  
; APPLICANT: O'Mahony, Daniel J.  
; TITLE OF INVENTION: LIPID-COMPRISING DRUG DELIVERY COMPLEXES  
; TITLE OF INVENTION: AND METHODS FOR THEIR PRODUCTION  
; FILE REFERENCE: 226272005300  
; CURRENT APPLICATION NUMBER: US/10/136,187  
; CURRENT FILING DATE: 2002-09-13  
; PRIOR APPLICATION NUMBER: US 60/287,786  
; PRIOR FILING DATE: 2001-04-30  
; NUMBER OF SEQ ID NOS: 45  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 39  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Synthetic Construct  
; LOCATION: 1  
; NAME/KEY: VARIANT  
; OTHER INFORMATION: Lysine is attached to DMPE-PEG5K-succinyl  
US-10-136-187-39

Query Match 100.0%; Score 68; DB 15; Length 15;  
Best Local Similarity 100.0%; Pred. No. 0.00029;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKAADVLLPVLLAAP 15  
|||||  
Db 1 KKAADVLLPVLLAAP 15

## RESULT 16

US-10-764-235-2  
; Sequence 2, Application US/10764235  
; Publication No. US20040138132A1  
; GENERAL INFORMATION:  
; APPLICANT: O'Mahony, Daniel J.  
; APPLICANT: Lambkin, Imelda J.

; TITLE OF INVENTION: MEMBRANE TRANSLOCATING PEPTIDE DRUG DELIVERY SYSTEM  
; FILE REFERENCE: P26,479-B USA  
; CURRENT APPLICATION NUMBER: US/10/764,235  
; CURRENT FILING DATE: 2004-01-23  
; PRIOR APPLICATION NUMBER: 09/671,089  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: 60/156,246  
; PRIOR FILING DATE: 1999-09-27  
; NUMBER OF SEQ ID NOS: 59  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 2  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: membrane translocating peptide  
; NAME/KEY: MOD RES  
; LOCATION: (15)..  
; OTHER INFORMATION: linked to FITC-LC  
US-10-764-235-2

Query Match 100.0%; Score 68; DB 16; Length 15;  
Best Local Similarity 100.0%; Pred. No. 0.00029;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKAADVLLPVLLAAP 15  
|||||  
Db 1 KKAADVLLPVLLAAP 15

## RESULT 17

US-10-850-873-2  
; Sequence 2, Application US/10850873  
; Publication No. US20050025821A1  
; GENERAL INFORMATION:  
; APPLICANT: Harvie, Pierrot  
; APPLICANT: Paul, Ralph  
; APPLICANT: Cudmore, Sally  
; APPLICANT: O'Mahony, Daniel J.  
; TITLE OF INVENTION: LIPID-COMPRISING DRUG DELIVERY COMPLEXES  
; TITLE OF INVENTION: AND METHODS FOR THEIR PRODUCTION  
; FILE REFERENCE: 226272005300  
; CURRENT APPLICATION NUMBER: US/10/850,873  
; CURRENT FILING DATE: 2004-05-20  
; PRIOR APPLICATION NUMBER: US/10/136,187  
; PRIOR FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: US 60/287,786  
; PRIOR FILING DATE: 2001-04-30  
; NUMBER OF SEQ ID NOS: 45  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Synthetic Construct  
US-10-850-873-2

Query Match 100.0%; Score 68; DB 17; Length 15;  
Best Local Similarity 100.0%; Pred. No. 0.00029;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKAADVLLPVLLAAP 15  
|||||  
Db 1 KKAADVLLPVLLAAP 15

## RESULT 18

US-10-850-873-6  
; Sequence 6, Application US/10850873  
; Publication No. US20050025821A1  
; GENERAL INFORMATION:

APPLICANT: Harvie, Pierrot  
APPLICANT: Paul, Ralph  
APPLICANT: Cudmore, Sally  
APPLICANT: O'Mahony, Daniel J.  
TITLE OF INVENTION: LIPID-COMPRISING DRUG DELIVERY COMPLEXES  
TITLE OF INVENTION: AND METHODS FOR THEIR PRODUCTION  
FILE REFERENCE: 226272005300  
CURRENT APPLICATION NUMBER: US/10/850,873  
CURRENT FILING DATE: 2004-05-20  
PRIOR APPLICATION NUMBER: US/10/136,187  
PRIOR FILING DATE: 2002-04-30  
PRIOR APPLICATION NUMBER: US 60/287,786  
PRIOR FILING DATE: 2001-04-30  
NUMBER OF SEQ ID NOS: 45  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 6  
LENGTH: 15  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Synthetic Construct  
NAME/KEY: VARIANT  
LOCATION: 1  
OTHER INFORMATION: Lysine is attached to galactose  
US-10-850-873-6

Query Match 100.0%; Score 68; DB 17; Length 15;  
Best Local Similarity 100.0%; Pred. No. 0.00029;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KKA0AVLLPVLLAAP 15  
Db 1 KKA0AVLLPVLLAAP 15

RESULT 19  
US-10-850-873-12  
Sequence 12, Application US/10850873  
Publication No. US20050025821A1  
GENERAL INFORMATION:  
APPLICANT: Harvie, Pierrot  
APPLICANT: Paul, Ralph  
APPLICANT: Cudmore, Sally  
APPLICANT: O'Mahony, Daniel J.  
TITLE OF INVENTION: LIPID-COMPRISING DRUG DELIVERY COMPLEXES  
TITLE OF INVENTION: AND METHODS FOR THEIR PRODUCTION  
FILE REFERENCE: 226272005300  
CURRENT APPLICATION NUMBER: US/10/850,873  
CURRENT FILING DATE: 2004-05-20  
PRIOR APPLICATION NUMBER: US/10/136,187  
PRIOR FILING DATE: 2002-04-30  
PRIOR APPLICATION NUMBER: US 60/287,786  
PRIOR FILING DATE: 2001-04-30  
NUMBER OF SEQ ID NOS: 45  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 12  
LENGTH: 15  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Synthetic Construct  
NAME/KEY: VARIANT  
LOCATION: 1  
OTHER INFORMATION: Lysine is attached to DOPE-succinyl  
US-10-850-873-12

Query Match 100.0%; Score 68; DB 17; Length 15;  
Best Local Similarity 100.0%; Pred. No. 0.00029;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KKA0AVLLPVLLAAP 15

Db 1 KKA0AVLLPVLLAAP 15  
RESULT 20  
US-10-850-873-13  
Sequence 13, Application US/10850873  
Publication No. US20050025821A1  
GENERAL INFORMATION:  
APPLICANT: Harvie, Pierrot  
APPLICANT: Paul, Ralph  
APPLICANT: Cudmore, Sally  
APPLICANT: O'Mahony, Daniel J.  
TITLE OF INVENTION: LIPID-COMPRISING DRUG DELIVERY COMPLEXES  
TITLE OF INVENTION: AND METHODS FOR THEIR PRODUCTION  
FILE REFERENCE: 226272005300  
CURRENT APPLICATION NUMBER: US/10/850,873  
CURRENT FILING DATE: 2004-05-20  
PRIOR APPLICATION NUMBER: US/10/136,187  
PRIOR FILING DATE: 2002-04-30  
PRIOR APPLICATION NUMBER: US 60/287,786  
PRIOR FILING DATE: 2001-04-30  
NUMBER OF SEQ ID NOS: 45  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 13  
LENGTH: 15  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Synthetic Construct  
US-10-850-873-13

Query Match 100.0%; Score 68; DB 17; Length 15;  
Best Local Similarity 100.0%; Pred. No. 0.00029;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KKA0AVLLPVLLAAP 15  
Db 1 KKA0AVLLPVLLAAP 15

RESULT 21  
US-10-850-873-33  
Sequence 33, Application US/10850873  
Publication No. US20050025821A1  
GENERAL INFORMATION:  
APPLICANT: Harvie, Pierrot  
APPLICANT: Paul, Ralph  
APPLICANT: Cudmore, Sally  
APPLICANT: O'Mahony, Daniel J.  
TITLE OF INVENTION: LIPID-COMPRISING DRUG DELIVERY COMPLEXES  
TITLE OF INVENTION: AND METHODS FOR THEIR PRODUCTION  
FILE REFERENCE: 226272005300  
CURRENT APPLICATION NUMBER: US/10/850,873  
CURRENT FILING DATE: 2004-05-20  
PRIOR APPLICATION NUMBER: US/10/136,187  
PRIOR FILING DATE: 2002-04-30  
PRIOR APPLICATION NUMBER: US 60/287,786  
PRIOR FILING DATE: 2001-04-30  
NUMBER OF SEQ ID NOS: 45  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 33  
LENGTH: 15  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Synthetic Construct  
NAME/KEY: VARIANT  
LOCATION: 1  
OTHER INFORMATION: Lysine is attached to Cholesteryl-succinyl  
US-10-850-873-33

```
Query Match      100.0%; Score 68; DB 17; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.00029;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKAADVLLPVLAAAP 15
DB 1 KKAADVLLPVLAAAP 15

RESULT 22
US-10-850-873-34
; Sequence 34, Application US/10850873
; Publication No. US20050025821A1
; GENERAL INFORMATION:
; APPLICANT: Harvie, Pierrot
; APPLICANT: Paul, Ralph
; APPLICANT: Cudmore, Sally
; APPLICANT: O'Mahony, Daniel J.
; TITLE OF INVENTION: LIPID-COMPRISING DRUG DELIVERY COMPLEXES
; TITLE OF INVENTION: AND METHODS FOR THEIR PRODUCTION
; FILE REFERENCE: 226272005300
; CURRENT APPLICATION NUMBER: US/10/850,873
; CURRENT FILING DATE: 2004-05-20
; PRIOR APPLICATION NUMBER: US/10/136,187
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/287,786
; PRIOR FILING DATE: 2001-04-30
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 34
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Construct
; NAME/KEY: VARIANT
; LOCATION: 1
; OTHER INFORMATION: Lysine is attached to DOPE-succinyl
US-10-850-873-34

Query Match      100.0%; Score 68; DB 17; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.00029;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKAADVLLPVLAAAP 15
DB 1 KKAADVLLPVLAAAP 15

RESULT 23
US-10-850-873-35
; Sequence 35, Application US/10850873
; Publication No. US20050025821A1
; GENERAL INFORMATION:
; APPLICANT: Harvie, Pierrot
; APPLICANT: Paul, Ralph
; APPLICANT: Cudmore, Sally
; APPLICANT: O'Mahony, Daniel J.
; TITLE OF INVENTION: LIPID-COMPRISING DRUG DELIVERY COMPLEXES
; TITLE OF INVENTION: AND METHODS FOR THEIR PRODUCTION
; FILE REFERENCE: 226272005300
; CURRENT APPLICATION NUMBER: US/10/850,873
; CURRENT FILING DATE: 2004-05-20
; PRIOR APPLICATION NUMBER: US/10/136,187
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/287,786
; PRIOR FILING DATE: 2001-04-30
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 35
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Construct
; NAME/KEY: VARIANT
; LOCATION: 1
; OTHER INFORMATION: Lysine is attached to DSPE-PEG5K-succinyl
US-10-850-873-35

Query Match      100.0%; Score 68; DB 17; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.00029;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKAADVLLPVLAAAP 15
DB 1 KKAADVLLPVLAAAP 15

RESULT 24
US-10-850-873-36
; Sequence 36, Application US/10850873
; Publication No. US20050025821A1
; GENERAL INFORMATION:
; APPLICANT: Harvie, Pierrot
; APPLICANT: Paul, Ralph
; APPLICANT: Cudmore, Sally
; APPLICANT: O'Mahony, Daniel J.
; TITLE OF INVENTION: LIPID-COMPRISING DRUG DELIVERY COMPLEXES
; TITLE OF INVENTION: AND METHODS FOR THEIR PRODUCTION
; FILE REFERENCE: 226272005300
; CURRENT APPLICATION NUMBER: US/10/850,873
; CURRENT FILING DATE: 2004-05-20
; PRIOR APPLICATION NUMBER: US/10/136,187
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/287,786
; PRIOR FILING DATE: 2001-04-30
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 36
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Construct
; NAME/KEY: VARIANT
; LOCATION: 1
; OTHER INFORMATION: Lysine is attached to DSPE-PEG5K-succinyl
US-10-850-873-36

Query Match      100.0%; Score 68; DB 17; Length 15;
Best Local Similarity 100.0%; Pred. No. 0.00029;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKAADVLLPVLAAAP 15
DB 1 KKAADVLLPVLAAAP 15

RESULT 25
US-10-850-873-37
; Sequence 37, Application US/10850873
; Publication No. US20050025821A1
; GENERAL INFORMATION:
; APPLICANT: Harvie, Pierrot
; APPLICANT: Paul, Ralph
; APPLICANT: Cudmore, Sally
; APPLICANT: O'Mahony, Daniel J.
; TITLE OF INVENTION: LIPID-COMPRISING DRUG DELIVERY COMPLEXES
```

US-10-850-873-37

Query Match 100.0%; Score 68; DB 17; Length 15;  
Best Local Similarity 100.0%; Pred. No. 0.00029;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KKAARAVLLPVLLAAP 15  
| | | | | | | | | | | | | | |  
Db 1 KKAARAVLLPVLLAAP 15

RESULT 26

US-10-850-873-38

Sequence 38, Application US/10850873  
Publication No. US20050025821A1  
GENERAL INFORMATION:  
APPLICANT: Harvie, Pierrot  
APPLICANT: Paul, Ralph  
APPLICANT: Cudmore, Sally  
APPLICANT: O'Mahony, Daniel J.  
TITLE OF INVENTION: LIPID-COMPRISING DRUG DELIVERY COMPLEXES  
FILE REFERENCE: 226272005300  
CURRENT APPLICATION NUMBER: US/10/850,873  
CURRENT FILING DATE: 2004-05-20  
PRIOR FILING DATE: 2002-04-30  
PRIOR APPLICATION NUMBER: US 60/287,786  
PRIOR FILING DATE: 2001-04-30  
NUMBER OF SEQ ID NOS: 45  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 37  
LENGTH: 15  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Synthetic Construct  
NAME/KEY: VARIANT  
LOCATION: 1  
OTHER INFORMATION: Lysine is attached to DMPE-PEG5K-succinyl

US-10-850-873-39

Query Match 100.0%; Score 68; DB 17; Length 15;  
Best Local Similarity 100.0%; Pred. No. 0.00029;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KKAARAVLLPVLLAAP 15  
| | | | | | | | | | | | | | |  
Db 1 KKAARAVLLPVLLAAP 15

RESULT 27

US-10-850-873-38

Sequence 38, Application US/10850873  
Publication No. US20050025821A1  
GENERAL INFORMATION:  
APPLICANT: Harvie, Pierrot  
APPLICANT: Paul, Ralph  
APPLICANT: Cudmore, Sally  
APPLICANT: O'Mahony, Daniel J.  
TITLE OF INVENTION: LIPID-COMPRISING DRUG DELIVERY COMPLEXES  
FILE REFERENCE: 226272005300  
CURRENT APPLICATION NUMBER: US/10/850,873  
CURRENT FILING DATE: 2004-05-20  
PRIOR FILING DATE: 2002-04-30  
PRIOR APPLICATION NUMBER: US 60/287,786  
PRIOR FILING DATE: 2001-04-30  
NUMBER OF SEQ ID NOS: 45  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 38  
LENGTH: 15  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Synthetic Construct  
NAME/KEY: VARIANT  
LOCATION: 1  
OTHER INFORMATION: Lysine is attached to DSPE-PEG5K-succinyl

US-10-850-873-39

Query Match 100.0%; Score 68; DB 17; Length 15;  
Best Local Similarity 100.0%; Pred. No. 0.00029;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KKAARAVLLPVLLAAP 15  
| | | | | | | | | | | | | | |  
Db 1 KKAARAVLLPVLLAAP 15

RESULT 28

US-10-955-656-2

Sequence 2, Application US/10955656  
Publication No. US20050101762A1  
GENERAL INFORMATION:  
APPLICANT: O'Mahony, Daniel J.  
APPLICANT: Lambkin, Imelda J.  
APPLICANT: Pinilla, Clemencia  
APPLICANT: Houghten, Richard  
TITLE OF INVENTION: CONJUGATES OF MEMBRANE TRANSLOCATING AGENTS AND PHARMACEUTICALLY  
FILE REFERENCE: P26,481-A USA  
CURRENT APPLICATION NUMBER: US/10/955,656  
CURRENT FILING DATE: 2004-09-30  
PRIOR APPLICATION NUMBER: 10/126,845  
PRIOR FILING DATE: 2002-04-19  
PRIOR APPLICATION NUMBER: 09/671,089  
PRIOR FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: 60/156,246  
PRIOR FILING DATE: 1999-09-27  
NUMBER OF SEQ ID NOS: 119  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 2  
LENGTH: 15  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: membrane translocating peptide

US-10-955-656-2

Query Match 100.0%; Score 68; DB 17; Length 15;  
Best Local Similarity 100.0%; Pred. No. 0.00029;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKAADVLLPVLLAAP 15  
 DB 1 KKAADVLLPVLLAAP 15

RESULT 29

US-10-955-656-3  
 ; Sequence 3, Application US/10955656  
 ; Publication No. US20050101762A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: O'Mahony, Daniel J.  
 ; APPLICANT: Lambkin, Imelda J.  
 ; APPLICANT: Pinilla, Clemencia  
 ; APPLICANT: Houghten, Richard  
 ; TITLE OF INVENTION: CONJUGATES OF MEMBRANE TRANSLOCATING AGENTS AND PHARMACEUTICALLY  
 ; FILE REFERENCE: P26.481-A USA  
 ; CURRENT APPLICATION NUMBER: US/10/955,656  
 ; PRIOR FILING DATE: 2004-09-30  
 ; PRIOR APPLICATION NUMBER: 10/126,845  
 ; PRIOR FILING DATE: 2002-04-19  
 ; PRIOR APPLICATION NUMBER: 09/671,089  
 ; PRIOR FILING DATE: 2000-09-27  
 ; PRIOR APPLICATION NUMBER: 60/156,246  
 ; PRIOR FILING DATE: 1999-09-27  
 ; NUMBER OF SEQ ID NOS: 119  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 3  
 ; LENGTH: 15  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: membrane translocating peptide  
 ; NAME/KEY: MOD RES  
 ; LOCATION: (15)..(15)  
 ; OTHER INFORMATION: linked to FITC-IC  
 US-10-955-656-3

Query Match 100.0%; Score 68; DB 17; Length 15;  
 Best Local Similarity 100.0%; Pred. No. 0.00029;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKAADVLLPVLLAAP 15  
 DB 1 KKAADVLLPVLLAAP 15

RESULT 30

US-10-955-656-60  
 ; Sequence 60, Application US/10955656  
 ; Publication No. US20050101762A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: O'Mahony, Daniel J.  
 ; APPLICANT: Lambkin, Imelda J.  
 ; APPLICANT: Pinilla, Clemencia  
 ; APPLICANT: Houghten, Richard  
 ; TITLE OF INVENTION: CONJUGATES OF MEMBRANE TRANSLOCATING AGENTS AND PHARMACEUTICALLY  
 ; FILE REFERENCE: P26.481-A USA  
 ; CURRENT APPLICATION NUMBER: US/10/955,656  
 ; PRIOR FILING DATE: 2004-09-30  
 ; PRIOR APPLICATION NUMBER: 10/126,845  
 ; PRIOR FILING DATE: 2002-04-19  
 ; PRIOR APPLICATION NUMBER: 09/671,089  
 ; PRIOR FILING DATE: 2000-09-27  
 ; PRIOR APPLICATION NUMBER: 60/156,246  
 ; PRIOR FILING DATE: 1999-09-27  
 ; NUMBER OF SEQ ID NOS: 119  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 60  
 ; LENGTH: 15

Query Match 100.0%; Score 68; DB 17; Length 15;  
 Best Local Similarity 100.0%; Pred. No. 0.00029;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: D form peptide  
 ; NAME/KEY: MISC FEATURE  
 ; LOCATION: (1)..(15)  
 ; OTHER INFORMATION: D form amino acid  
 US-10-955-656-60

Query Match 100.0%; Score 68; DB 17; Length 15;  
 Best Local Similarity 100.0%; Pred. No. 0.00029;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKAADVLLPVLLAAP 15  
 DB 1 KKAADVLLPVLLAAP 15

RESULT 31

US-10-955-656-61  
 ; Sequence 61, Application US/10955656  
 ; Publication No. US20050101762A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: O'Mahony, Daniel J.  
 ; APPLICANT: Lambkin, Imelda J.  
 ; APPLICANT: Pinilla, Clemencia  
 ; APPLICANT: Houghten, Richard  
 ; TITLE OF INVENTION: CONJUGATES OF MEMBRANE TRANSLOCATING AGENTS AND PHARMACEUTICALLY  
 ; FILE REFERENCE: P26.481-A USA  
 ; CURRENT APPLICATION NUMBER: US/10/955,656  
 ; PRIOR FILING DATE: 2004-09-30  
 ; PRIOR APPLICATION NUMBER: 10/126,845  
 ; PRIOR FILING DATE: 2002-04-19  
 ; PRIOR APPLICATION NUMBER: 09/671,089  
 ; PRIOR FILING DATE: 2000-09-27  
 ; PRIOR APPLICATION NUMBER: 60/156,246  
 ; PRIOR FILING DATE: 1999-09-27  
 ; NUMBER OF SEQ ID NOS: 119  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 61  
 ; LENGTH: 15  
 ; TYPE: PRT  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: D form peptide  
 ; NAME/KEY: MISC FEATURE  
 ; LOCATION: (1)..(15)  
 ; OTHER INFORMATION: D form amino acid  
 US-10-955-656-61

Query Match 100.0%; Score 68; DB 17; Length 15;  
 Best Local Similarity 100.0%; Pred. No. 0.00029;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKAADVLLPVLLAAP 15  
 DB 1 KKAADVLLPVLLAAP 15

RESULT 32

US-10-126-845-48  
 ; Sequence 48, Application US/10126845  
 ; Publication No. US20030181367A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: O'Mahony, Daniel J.  
 ; APPLICANT: Lambkin, Imelda J.  
 ; APPLICANT: Pinilla, Clemencia  
 ; APPLICANT: Houghten, Richard  
 ; TITLE OF INVENTION: MEMBRANE TRANSLOCATING PEPTIDE DRUG DELIVERY SYSTEM  
 ; FILE REFERENCE: E1067/20058

Query Match 100.0%; Score 68; DB 17; Length 15;  
 Best Local Similarity 100.0%; Pred. No. 0.00029;  
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;



; CURRENT APPLICATION NUMBER: US/10/126,845  
; CURRENT FILING DATE: 2002-10-15  
; NUMBER OF SEQ ID NOS: 119  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 48  
; LENGTH: 16  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: dansylated membrane translocating peptide  
; NAME/KEY: MOD.RES  
; LOCATION: (1)..(1)  
; OTHER INFORMATION: dansylated  
US-10-126-845-48

Query Match 100.0%; Score 68; DB 14; Length 16;  
Best Local Similarity 100.0%; Pred. No. 0.00031;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 KKAQAVLLPVLLAAP 15  
Db 2 KKAQAVLLPVLLAAP 16

RESULT 33  
US-10-136-187-3  
; Sequence 3, Application US/10136187  
; Publication No. US20030203865A1  
; GENERAL INFORMATION:  
; APPLICANT: Harvie, Pierrot  
; APPLICANT: Paul, Ralph  
; APPLICANT: Cudmore, Sally  
; APPLICANT: O'Mahony, Daniel J.  
; TITLE OF INVENTION: LIPID-COMPRISING DRUG DELIVERY COMPLEXES  
; TITLE OF INVENTION: AND METHODS FOR THEIR PRODUCTION  
; FILE REFERENCE: 226272005300  
; CURRENT APPLICATION NUMBER: US/10/136,187  
; CURRENT FILING DATE: 2002-09-13  
; PRIOR APPLICATION NUMBER: US 60/287,786  
; PRIOR FILING DATE: 2001-04-30  
; NUMBER OF SEQ ID NOS: 45  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 16  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Synthetic Construct  
US-10-136-187-3

Query Match 100.0%; Score 68; DB 15; Length 16;  
Best Local Similarity 100.0%; Pred. No. 0.00031;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 KKAQAVLLPVLLAAP 15  
Db 2 KKAQAVLLPVLLAAP 16

RESULT 34  
US-10-136-187-7  
; Sequence 7, Application US/10136187  
; Publication No. US20030203865A1  
; GENERAL INFORMATION:  
; APPLICANT: Harvie, Pierrot  
; APPLICANT: Paul, Ralph  
; APPLICANT: Cudmore, Sally  
; APPLICANT: O'Mahony, Daniel J.  
; TITLE OF INVENTION: LIPID-COMPRISING DRUG DELIVERY COMPLEXES  
; TITLE OF INVENTION: AND METHODS FOR THEIR PRODUCTION  
; FILE REFERENCE: 226272005300  
; CURRENT APPLICATION NUMBER: US/10/136,187

; CURRENT FILING DATE: 2002-09-13  
; PRIOR APPLICATION NUMBER: US 60/287,786  
; PRIOR FILING DATE: 2001-04-30  
; NUMBER OF SEQ ID NOS: 45  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 7  
; LENGTH: 16  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Synthetic Construct  
; NAME/KEY: VARIANT  
; LOCATION: 1  
; OTHER INFORMATION: Lysine is attached to galactose  
US-10-136-187-7

Query Match 100.0%; Score 68; DB 15; Length 16;  
Best Local Similarity 100.0%; Pred. No. 0.00031;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 KKAQAVLLPVLLAAP 15  
Db 2 KKAQAVLLPVLLAAP 16

RESULT 35  
US-10-136-187-28  
; Sequence 28, Application US/10136187  
; Publication No. US20030203865A1  
; GENERAL INFORMATION:  
; APPLICANT: Harvie, Pierrot  
; APPLICANT: Paul, Ralph  
; APPLICANT: Cudmore, Sally  
; APPLICANT: O'Mahony, Daniel J.  
; TITLE OF INVENTION: LIPID-COMPRISING DRUG DELIVERY COMPLEXES  
; TITLE OF INVENTION: AND METHODS FOR THEIR PRODUCTION  
; FILE REFERENCE: 226272005300  
; CURRENT APPLICATION NUMBER: US/10/136,187  
; CURRENT FILING DATE: 2002-09-13  
; PRIOR APPLICATION NUMBER: US 60/287,786  
; PRIOR FILING DATE: 2001-04-30  
; NUMBER OF SEQ ID NOS: 45  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 28  
; LENGTH: 16  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Synthetic Construct  
; NAME/KEY: VARIANT  
; LOCATION: 1  
; OTHER INFORMATION: Lysine is attached to dansyl  
US-10-136-187-28

Query Match 100.0%; Score 68; DB 15; Length 16;  
Best Local Similarity 100.0%; Pred. No. 0.00031;  
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 KKAQAVLLPVLLAAP 15  
Db 2 KKAQAVLLPVLLAAP 16

RESULT 36  
US-10-136-187-32  
; Sequence 32, Application US/10136187  
; Publication No. US20030203865A1  
; GENERAL INFORMATION:  
; APPLICANT: Harvie, Pierrot  
; APPLICANT: Paul, Ralph  
; APPLICANT: Cudmore, Sally



```
; APPLICANT: Cudmore, Sally
; APPLICANT: O'Mahony, Daniel J.
; TITLE OF INVENTION: LIPID-COMPRISING DRUG DELIVERY COMPLEXES
; FILE REFERENCE: 226272005300
; CURRENT APPLICATION NUMBER: US/10/850,873
; CURRENT FILING DATE: 2004-05-20
; PRIOR APPLICATION NUMBER: US/10/136,187
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/287,786
; PRIOR FILING DATE: 2001-04-30
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Construct
US-10-850-873-3
```

```
Query Match 100.0%; Score 68; DB 17; Length 16;
Best Local Similarity 100.0%; Pred. No. 0.00031;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 KKAARVLLPVLLAAP 15
    |||||
Db 2 KKAARVLLPVLLAAP 16
```

```
RESULT 41
US-10-850-873-7
; Sequence 7, Application US/10850873
; Publication No. US20050025821A1
; GENERAL INFORMATION:
; APPLICANT: Harvie, Pierrot
; APPLICANT: Paul, Ralph
; APPLICANT: Cudmore, Sally
; APPLICANT: O'Mahony, Daniel J.
; TITLE OF INVENTION: LIPID-COMPRISING DRUG DELIVERY COMPLEXES
; FILE REFERENCE: 226272005300
; CURRENT APPLICATION NUMBER: US/10/850,873
; CURRENT FILING DATE: 2004-05-20
; PRIOR APPLICATION NUMBER: US/10/136,187
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/287,786
; PRIOR FILING DATE: 2001-04-30
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Construct
; NAME/KEY: VARIANT
; LOCATION: 1
; OTHER INFORMATION: Lysine is attached to galactose
US-10-850-873-7
```

```
Query Match 100.0%; Score 68; DB 17; Length 16;
Best Local Similarity 100.0%; Pred. No. 0.00031;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 KKAARVLLPVLLAAP 15
    |||||
Db 2 KKAARVLLPVLLAAP 16
```

```
RESULT 42
US-10-850-873-28
```

```
; Sequence 28, Application US/10850873
; Publication No. US20050025821A1
; GENERAL INFORMATION:
; APPLICANT: Harvie, Pierrot
; APPLICANT: Paul, Ralph
; APPLICANT: Cudmore, Sally
; APPLICANT: O'Mahony, Daniel J.
; TITLE OF INVENTION: LIPID-COMPRISING DRUG DELIVERY COMPLEXES
; FILE REFERENCE: 226272005300
; CURRENT APPLICATION NUMBER: US/10/850,873
; CURRENT FILING DATE: 2004-05-20
; PRIOR APPLICATION NUMBER: US/10/136,187
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/287,786
; PRIOR FILING DATE: 2001-04-30
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 28
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Construct
; NAME/KEY: VARIANT
; LOCATION: 1
; OTHER INFORMATION: Lysine is attached to dansyl
US-10-850-873-28
```

```
Query Match 100.0%; Score 68; DB 17; Length 16;
Best Local Similarity 100.0%; Pred. No. 0.00031;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 1 KKAARVLLPVLLAAP 15
    |||||
Db 2 KKAARVLLPVLLAAP 16
```

```
RESULT 43
US-10-850-873-32
; Sequence 32, Application US/10850873
; Publication No. US20050025821A1
; GENERAL INFORMATION:
; APPLICANT: Harvie, Pierrot
; APPLICANT: Paul, Ralph
; APPLICANT: Cudmore, Sally
; APPLICANT: O'Mahony, Daniel J.
; TITLE OF INVENTION: LIPID-COMPRISING DRUG DELIVERY COMPLEXES
; FILE REFERENCE: 226272005300
; CURRENT APPLICATION NUMBER: US/10/850,873
; CURRENT FILING DATE: 2004-05-20
; PRIOR APPLICATION NUMBER: US/10/136,187
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/287,786
; PRIOR FILING DATE: 2001-04-30
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 32
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Construct
; NAME/KEY: VARIANT
; LOCATION: 1
; OTHER INFORMATION: Serine is attached to galactose
US-10-850-873-32
```

```
Query Match 100.0%; Score 68; DB 17; Length 16;
Best Local Similarity 100.0%; Pred. No. 0.00031;
```

```
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKAADVLLPVLLAAP 15
    |||||
Db 2 KKAADVLLPVLLAAP 16
    |||||

RESULT 44
US-10-955-656-48
; Sequence 48, Application US/109555656
; Publication No. US20050101762A1
; GENERAL INFORMATION:
; APPLICANT: O'Mahony, Daniel J.
; APPLICANT: Lambkin, Imelda J.
; APPLICANT: Pinilla, Clemencia
; APPLICANT: Houghten, Richard
; TITLE OF INVENTION: CONJUGATES OF MEMBRANE TRANSLOCATING AGENTS AND PHARMACEUTICALLY
; FILE REFERENCE: P26,481-A USA
; CURRENT APPLICATION NUMBER: US/10/955,656
; CURRENT FILING DATE: 2004-09-30
; PRIOR APPLICATION NUMBER: 10/126,845
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 09/671,089
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/156,246
; PRIOR FILING DATE: 1999-09-27
; NUMBER OF SEQ ID NOS: 119
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 48
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: dansylated membrane translocating peptide
; NAME/KEY: MOD RES
; LOCATION: (1)-(1)
; OTHER INFORMATION: dansylated
US-10-955-656-48

Query Match 100.0%; Score 68; DB 17; Length 16;
Best Local Similarity 100.0%; Pred. No. 0.00031;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKAADVLLPVLLAAP 15
    |||||
Db 2 KKAADVLLPVLLAAP 16
    |||||

RESULT 45
US-10-136-187-30
; Sequence 30, Application US/10136187
; Publication No. US20030203865A1
; GENERAL INFORMATION:
; APPLICANT: Harvie, Pierrot
; APPLICANT: Paul, Ralph
; APPLICANT: O'Mahony, Daniel J.
; APPLICANT: Cudmore, Sally
; TITLE OF INVENTION: LIPID-COMPRISING DRUG DELIVERY COMPLEXES
; FILE REFERENCE: 226272005300
; CURRENT APPLICATION NUMBER: US/10/136,187
; CURRENT FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: US 60/287,786
; PRIOR FILING DATE: 2001-04-30
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 30
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: dansylated peptide translocating
```

```
; OTHER INFORMATION: Synthetic Construct
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 1
; OTHER INFORMATION: Lysine is attached to dansyl
US-10-136-187-30

Query Match 100.0%; Score 68; DB 15; Length 17;
Best Local Similarity 100.0%; Pred. No. 0.00033;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKAADVLLPVLLAAP 15
    |||||
Db 3 KKAADVLLPVLLAAP 17
    |||||

RESULT 46
US-10-850-873-30
; Sequence 30, Application US/10850873
; Publication No. US20050025821A1
; GENERAL INFORMATION:
; APPLICANT: Harvie, Pierrot
; APPLICANT: Paul, Ralph
; APPLICANT: Cudmore, Sally
; APPLICANT: O'Mahony, Daniel J.
; TITLE OF INVENTION: LIPID-COMPRISING DRUG DELIVERY COMPLEXES
; FILE REFERENCE: 226272005300
; CURRENT APPLICATION NUMBER: US/10/850,873
; CURRENT FILING DATE: 2004-05-20
; PRIOR APPLICATION NUMBER: US/10/136,187
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/287,786
; PRIOR FILING DATE: 2001-04-30
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 30
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Construct
; NAME/KEY: VARIANT
; LOCATION: 1
; OTHER INFORMATION: Lysine is attached to dansyl
US-10-850-873-30

Query Match 100.0%; Score 68; DB 17; Length 17;
Best Local Similarity 100.0%; Pred. No. 0.00033;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKAADVLLPVLLAAP 15
    |||||
Db 3 KKAADVLLPVLLAAP 17
    |||||

RESULT 47
US-10-126-845-4
; Sequence 4, Application US/10126845
; Publication No. US20030181367A1
; GENERAL INFORMATION:
; APPLICANT: O'Mahony, Daniel J.
; APPLICANT: Lambkin, Imelda J.
; APPLICANT: Pinilla, Clemencia
; APPLICANT: Houghten, Richard
; TITLE OF INVENTION: MEMBRANE TRANSLOCATING PEPTIDE DRUG DELIVERY SYSTEM
; FILE REFERENCE: E1067/20058
; CURRENT APPLICATION NUMBER: US/10/126,845
; CURRENT FILING DATE: 2002-10-15
; NUMBER OF SEQ ID NOS: 119
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
```

```
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: membrane translocating peptide
US-10-126-845-4

Query Match          100.0%; Score 68; DB 14; Length 19;
Best Local Similarity 100.0%; Pred. No. 0.00037;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KKAARVLLPVLLAAP 15
Db 1 KKAARVLLPVLLAAP 15

RESULT 48
US-10-126-845-62
; Sequence 62, Application US/10126845
; Publication No. US20030181367A1
; GENERAL INFORMATION:
; APPLICANT: O'Mahony, Daniel J.
; APPLICANT: Lambkin, Imelda J.
; APPLICANT: Pinilla, Clemencia
; APPLICANT: Houghten, Richard
; TITLE OF INVENTION: MEMBRANE TRANSLOCATING PEPTIDE DRUG DELIVERY SYSTEM
; FILE REFERENCE: E1067/20058
; CURRENT APPLICATION NUMBER: US/10/126,845
; CURRENT FILING DATE: 2002-10-15
; NUMBER OF SEQ ID NOS: 119
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 62
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: D form peptide
; NAME/KEY: MISC FEATURE
; LOCATION: (1)-(19)
; OTHER INFORMATION: D form amino acid
US-10-126-845-62

Query Match          100.0%; Score 68; DB 14; Length 19;
Best Local Similarity 100.0%; Pred. No. 0.00037;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KKAARVLLPVLLAAP 15
Db 1 KKAARVLLPVLLAAP 15

RESULT 49
US-10-126-845-108
; Sequence 108, Application US/10126845
; Publication No. US20030181367A1
; GENERAL INFORMATION:
; APPLICANT: O'Mahony, Daniel J.
; APPLICANT: Lambkin, Imelda J.
; APPLICANT: Pinilla, Clemencia
; APPLICANT: Houghten, Richard
; TITLE OF INVENTION: MEMBRANE TRANSLOCATING PEPTIDE DRUG DELIVERY SYSTEM
; FILE REFERENCE: E1067/20058
; CURRENT APPLICATION NUMBER: US/10/126,845
; CURRENT FILING DATE: 2002-10-15
; NUMBER OF SEQ ID NOS: 119
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 108
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: D-form peptide; Comprises opioid peptide
```

```
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (1)-(19)
; OTHER INFORMATION: D-form amino acid
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (19)..(19)
; OTHER INFORMATION: Amidation
US-10-126-845-108

Query Match          100.0%; Score 68; DB 14; Length 19;
Best Local Similarity 100.0%; Pred. No. 0.00037;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KKAARVLLPVLLAAP 15
Db 5 KKAARVLLPVLLAAP 19

RESULT 50
US-10-116-275-92
; Sequence 92, Application US/10116275
; Publication No. US20030211476A1
; GENERAL INFORMATION:
; APPLICANT: Elan Pharmaceutical Technology
; APPLICANT: O'Mahony, Daniel J.
; APPLICANT: Brayden, David
; APPLICANT: Byrne, Daragh
; APPLICANT: Lambkin, Imelda
; APPLICANT: Higgins, Lisa
; TITLE OF INVENTION: Genetic Analysis of Peyer's Patches and M Cells and Methods and
; FILE REFERENCE: E1067/20087
; CURRENT APPLICATION NUMBER: US/10/116,275
; CURRENT FILING DATE: 2002-10-04
; NUMBER OF SEQ ID NOS: 349
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 92
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Listed in Table titled "Peptides that Target to and/or Enhance Up
; OTHER INFORMATION: take Across the GIT"
US-10-116-275-92

Query Match          100.0%; Score 68; DB 15; Length 19;
Best Local Similarity 100.0%; Pred. No. 0.00037;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KKAARVLLPVLLAAP 15
Db 1 KKAARVLLPVLLAAP 15

Search completed: June 2, 2005, 01:43:48
Job time : 89.6562 secs
```

THIS PAGE BLANK (07/25/00)

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 2, 2005, 01:27:36 ; Search time 31.875 Seconds  
(without alignments)  
39.813 Million cell updates/sec

Title: US-10-764-235-24

Perfect score: 76  
Sequence: 1 KKKRKAATAVLLPVLIA 17

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 750 summaries

Database : Issued Patents AA:\*

1: /cgn2\_6/prodata/1/iaa/5A COMB.pep.\*  
2: /cgn2\_6/prodata/1/iaa/5B COMB.pep.\*  
3: /cgn2\_6/prodata/1/iaa/6A COMB.pep.\*  
4: /cgn2\_6/prodata/1/iaa/6B COMB.pep.\*  
5: /cgn2\_6/prodata/1/iaa/PCTUS COMB.pep.\*  
6: /cgn2\_6/prodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	76	100.0	17	4	US-09-671-089-24
2	53.5	70.4	29	2	US-08-928-958-24
3	53.5	70.4	29	2	US-09-072-429-24
4	53.5	70.4	30	2	US-08-928-958-1
5	53.5	70.4	30	2	US-09-072-429-1
6	53	69.7	16	4	US-09-671-089-3
7	53	69.7	16	4	US-09-671-089-48
8	51	67.1	15	4	US-09-671-089-2
9	51	67.1	19	4	US-09-671-089-4
10	47	61.8	14	4	US-09-671-089-7
11	47	61.8	14	4	US-09-671-089-58
12	47	61.8	16	4	US-09-671-089-5
13	47	61.8	26	2	US-08-928-958-2
14	47	61.8	26	2	US-09-072-429-2
15	44	57.9	19	4	US-10-116-288A-22
16	44	57.9	25	4	US-10-116-288A-20
17	43	56.6	10	3	US-09-186-170-4
18	43	56.6	10	4	US-09-562-868-4
19	43	56.6	10	4	US-10-116-288A-4
20	43	56.6	11	3	US-09-186-170-5
21	43	56.6	11	4	US-09-562-868-5
22	43	56.6	11	4	US-10-116-288A-5
23	43	56.6	11	4	US-09-671-089-15
24	43	56.6	12	3	US-09-186-170-1
25	43	56.6	12	4	US-09-562-868-1
26	43	56.6	12	4	US-09-997-465B-4
27	43	56.6	12	4	US-10-083-889-8

28	43	56.6	12	4	US-10-116-288A-1	Sequence 1, Appli
29	43	56.6	12	4	US-09-671-089-1	Sequence 1, Appli
30	43	56.6	12	4	US-09-671-089-14	Sequence 14, Appli
31	43	56.6	12	4	US-09-671-089-16	Sequence 16, Appli
32	43	56.6	12	4	US-09-671-089-59	Sequence 59, Appli
33	43	56.6	12	4	US-10-144-549-19	Sequence 6, Appli
34	43	56.6	13	4	US-09-671-089-6	Sequence 23, Appli
35	43	56.6	17	4	US-09-671-089-23	Sequence 20502, A
36	43	56.6	456	4	US-09-252-991A-20502	Sequence 7, Appli
37	41	53.9	30	3	US-08-448-489-7	Sequence 7, Appli
38	41	53.9	30	4	US-09-689-730-7	Sequence 19092, A
39	41	53.9	421	4	US-09-252-991A-19092	Sequence 1, Appli
40	41	53.9	579	3	US-08-704-711A-1	Sequence 1, Appli
41	41	53.9	579	3	US-09-521-220-1	Sequence 2, Appli
42	41	53.9	582	3	US-08-704-711A-2	Sequence 1, Appli
43	41	53.9	582	3	US-08-448-489-1	Sequence 1, Appli
44	41	53.9	582	3	US-09-211-704A-9	Sequence 9, Appli
45	41	53.9	582	3	US-09-521-220-2	Sequence 2, Appli
46	41	53.9	582	3	US-09-391-104-28	Sequence 84, Appli
47	41	53.9	582	4	US-09-919-497-88	Sequence 1, Appli
48	41	53.9	582	4	US-09-689-730-1	Sequence 41668, A
49	40.5	53.3	716	4	US-09-270-767-41668	Sequence 20037, A
50	40	52.6	238	4	US-09-252-991A-20037	Sequence 2954, Ap
51	40	52.6	513	4	US-09-710-279-2954	Sequence 5001, Ap
52	40	52.6	551	3	US-09-134-001C-5001	Sequence 3, Appli
53	39	51.3	9	3	US-09-186-170-3	Sequence 3, Appli
54	39	51.3	9	4	US-09-562-868-3	Sequence 3, Appli
55	39	51.3	9	4	US-10-116-288A-3	Sequence 17, Appli
56	39	51.3	9	4	US-09-671-089-17	Sequence 9, Appli
57	39	51.3	11	3	US-09-186-170-9	Sequence 9, Appli
58	39	51.3	11	4	US-09-562-868-9	Sequence 9, Appli
59	39	51.3	11	4	US-10-116-288A-9	Sequence 8, Appli
60	39	51.3	11	4	US-09-671-089-8	Sequence 10, Appli
61	39	51.3	11	4	US-09-671-089-19	Sequence 10, Appli
62	39	51.3	13	4	US-09-671-089-10	Sequence 10, Appli
63	39	51.3	206	4	US-09-270-767-46825	Sequence 46825, A
64	39	51.3	248	4	US-09-107-532A-6599	Sequence 6599, Ap
65	39	51.3	260	4	US-09-902-540-14188	Sequence 14188, A
66	39	51.3	329	4	US-09-252-991A-18420	Sequence 18420, A
67	39	51.3	344	4	US-09-270-767-41438	Sequence 41438, A
68	39	51.3	644	4	US-09-252-991A-25084	Sequence 25084, A
69	38	50.0	14	2	US-08-503-226B-41	Sequence 41, Appli
70	38	50.0	14	3	US-08-721-458B-41	Sequence 19, Appli
71	38	50.0	28	3	US-09-348-578-19	Sequence 19, Appli
72	38	50.0	28	4	US-09-699-684-19	Sequence 19, Appli
73	38	50.0	71	4	US-09-732-210-108	Sequence 108, App
74	38	50.0	135	4	US-09-252-991A-17432	Sequence 17432, A
75	38	50.0	232	4	US-09-107-532A-6091	Sequence 6091, Ap
76	38	50.0	296	4	US-09-540-236-2856	Sequence 2856, Ap
77	38	50.0	309	4	US-09-522-714-10	Sequence 10, Appli
78	38	50.0	418	4	US-09-966-921A-2	Sequence 2, Appli
79	38	50.0	467	4	US-09-107-532A-6175	Sequence 6175, Ap
80	38	50.0	483	4	US-09-902-540-9804	Sequence 9804, Ap
81	38	50.0	515	1	US-10-095-946-10	Sequence 10, Appli
82	38	50.0	515	3	US-09-183-959-10	Sequence 10, Appli
83	38	50.0	515	4	US-09-347-650-10	Sequence 10, Appli
84	38	50.0	515	4	US-09-535-315-10	Sequence 10, Appli
85	38	50.0	1114	4	US-09-252-991A-24965	Sequence 24965, A
86	37	48.7	179	4	US-09-902-540-9821	Sequence 9821, Ap
87	37	48.7	191	4	US-09-543-681A-8048	Sequence 8048, Ap
88	37	48.7	279	3	US-09-097-767A-10	Sequence 10, Appli
89	37	48.7	279	3	US-09-097-767A-14	Sequence 14, Appli
90	37	48.7	287	3	US-09-097-767A-17	Sequence 17, Appli
91	37	48.7	354	1	US-08-313-553-9	Sequence 9, Appli
92	37	48.7	354	3	US-08-767-993-9	Sequence 23, Appli
93	37	48.7	387	3	US-09-097-767A-23	Sequence 23, Appli
94	37	48.7	421	4	US-09-252-991A-26478	Sequence 26478, A
95	37	48.7	562	4	US-08-851-567B-30	Sequence 30, Appli
96	37	48.7	1189	4	US-08-851-567B-26	Sequence 26, Appli
97	36.5	48.0	27	2	US-08-928-958-3	Sequence 3, Appli
98	36.5	48.0	27	2	US-09-072-429-3	Sequence 3, Appli
99	36	47.4	73	4	US-09-590-968B-17	Sequence 17, Appli
100	36	47.4	126	4	US-09-489-039A-11199	Sequence 11199, A

101	36	47.4	165	4	US-09-902-540-10298	Sequence 10298, A	174	35	46.1	805	4	US-09-543-681A-7900	Sequence 7900, Ap
102	36	47.4	135	4	US-09-302-540-11422	Sequence 11422, A	175	35	46.1	822	2	US-08-474-067-9	Sequence 9, Appli
103	36	47.4	208	4	US-09-569-611C-27	Sequence 27, Appl	176	35	46.1	822	2	US-08-474-068A-9	Sequence 9, Appli
104	36	47.4	210	4	US-09-569-611C-28	Sequence 28, Appl	177	35	46.1	822	2	US-08-472-481-8	Sequence 8, Appli
105	36	47.4	304	4	US-10-000-489-80	Sequence 80, Appl	178	34.5	45.4	283	4	US-09-248-796A-16847	Sequence 16847, A
106	36	47.4	307	4	US-09-893-737-304	Sequence 304, App	179	34.5	45.4	1037	4	US-09-252-991A-17548	Sequence 17548, A
107	36	47.4	328	4	US-09-902-540-10796	Sequence 10796, A	180	34	44.7	50	4	US-09-621-976-6033	Sequence 6033, Ap
108	36	47.4	352	4	US-09-270-767-42111	Sequence 42211, A	181	34	44.7	61	4	US-09-248-796A-21045	Sequence 21045, A
109	36	47.4	415	4	US-09-543-681A-4986	Sequence 4986, Ap	182	34	44.7	62	2	US-08-530-569B-21	Sequence 21, Appl
110	36	47.4	415	4	US-09-252-991A-27213	Sequence 27213, A	183	34	44.7	67	4	US-09-621-976-5505	Sequence 5505, Ap
111	36	47.4	445	4	US-09-902-540-13038	Sequence 13038, A	184	34	44.7	71	4	US-09-248-796A-26716	Sequence 26716, A
112	36	47.4	445	4	US-09-902-540-13038	Sequence 13038, A	184	34	44.7	71	4	US-09-248-796A-26716	Sequence 26716, A
113	36	47.4	451	4	US-09-949-016-7978	Sequence 7978, Ap	185	34	44.7	82	4	US-09-107-532A-6203	Sequence 6203, Ap
114	36	47.4	451	4	US-09-949-016-7979	Sequence 7979, Ap	186	34	44.7	82	4	US-09-270-767-46251	Sequence 46251, A
115	36	47.4	451	4	US-09-949-016-7979	Sequence 7979, Ap	186	34	44.7	82	4	US-09-270-767-46251	Sequence 46251, A
116	36	47.4	451	4	US-09-949-016-7981	Sequence 7980, Ap	187	34	44.7	86	4	US-09-248-796A-24870	Sequence 24870, A
117	36	47.4	474	4	US-09-949-016-7982	Sequence 7981, Ap	188	34	44.7	117	4	US-09-248-796A-18571	Sequence 18571, A
118	36	47.4	474	4	US-09-949-016-7983	Sequence 7982, Ap	188	34	44.7	117	4	US-09-248-796A-18571	Sequence 18571, A
119	36	47.4	474	4	US-09-949-016-7985	Sequence 7985, Ap	189	34	44.7	126	4	US-09-489-039A-8610	Sequence 8610, Ap
120	36	47.4	474	4	US-09-949-016-7986	Sequence 7986, Ap	189	34	44.7	126	4	US-09-489-039A-8610	Sequence 8610, Ap
121	36	47.4	474	4	US-09-949-016-7987	Sequence 7987, Ap	189	34	44.7	126	4	US-09-489-039A-8610	Sequence 8610, Ap
122	36	47.4	578	4	US-09-252-991A-18288	Sequence 18288, A	190	34	44.7	133	4	US-09-902-540-14536	Sequence 14536, A
123	36	47.4	592	4	US-08-946-412-2	Sequence 2, Appli	191	34	44.7	143	4	US-09-270-767-35980	Sequence 35980, A
124	36	47.4	823	3	US-09-651-656-19	Sequence 19, Appl	192	34	44.7	143	4	US-09-270-767-35980	Sequence 35980, A
125	36	47.4	823	3	US-09-651-656-19	Sequence 19, Appl	192	34	44.7	143	4	US-09-270-767-35980	Sequence 35980, A
126	36	47.4	1722	4	US-09-538-092-1033	Sequence 1033, Ap	193	34	44.7	165	4	US-09-489-039A-11784	Sequence 11784, A
127	36	47.4	1909	4	US-09-590-968B-2	Sequence 2, Appli	194	34	44.7	220	4	US-09-270-767-36442	Sequence 36442, A
128	35	46.1	4	3	US-09-186-170-2	Sequence 2, Appli	200	34	44.7	329	4	US-09-602-787A-306	Sequence 306, App
129	35	46.1	8	4	US-09-562-868-2	Sequence 2, Appli	201	34	44.7	332	4	US-09-270-767-42655	Sequence 42655, A
130	35	46.1	8	4	US-10-116-288A-2	Sequence 2, Appli	202	34	44.7	346	4	US-09-170-496D-254	Sequence 254, App
131	35	46.1	8	4	US-09-671-089-18	Sequence 18, Appl	203	34	44.7	346	4	US-09-170-496D-254	Sequence 254, App
132	35	46.1	10	3	US-09-186-170-8	Sequence 8, Appli	204	34	44.7	366	4	US-09-489-039A-13929	Sequence 12929, A
133	35	46.1	10	4	US-09-562-868-8	Sequence 8, Appli	205	34	44.7	377	4	US-09-910-430-27	Sequence 27, Appl
134	35	46.1	10	4	US-09-671-089-9	Sequence 9, Appli	206	34	44.7	399	4	US-09-489-039A-7814	Sequence 7814, Ap
135	35	46.1	10	4	US-09-671-089-20	Sequence 20, Appl	207	34	44.7	401	2	US-08-820-521-2	Sequence 2, Appli
136	35	46.1	12	4	US-09-671-089-11	Sequence 11, Appl	208	34	44.7	401	3	US-09-248-715-2	Sequence 2, Appli
137	35	46.1	12	4	US-09-671-089-11	Sequence 11, Appl	209	34	44.7	401	3	US-09-248-715-2	Sequence 2, Appli
138	35	46.1	36	4	US-09-270-767-56911	Sequence 56911, A	210	34	44.7	410	4	US-09-949-016-10345	Sequence 10345, A
139	35	46.1	40	4	US-09-902-540-14864	Sequence 14864, A	211	34	44.7	410	4	US-09-949-016-10345	Sequence 10345, A
140	35	46.1	91	4	US-09-071-035-382	Sequence 382, App	212	34	44.7	459	4	US-09-949-016-10346	Sequence 10346, A
141	35	46.1	91	4	US-09-071-035-382	Sequence 382, App	213	34	44.7	459	4	US-09-949-016-10346	Sequence 10346, A
142	35	46.1	109	4	US-09-248-796A-27219	Sequence 27219, A	214	34	44.7	473	4	US-09-270-767-43452	Sequence 43452, A
143	35	46.1	123	4	US-09-248-796A-24006	Sequence 24006, A	215	34	44.7	473	4	US-09-270-767-43452	Sequence 43452, A
144	35	46.1	123	4	US-09-134-000C-5564	Sequence 5564, Ap	216	34	44.7	473	4	US-09-270-767-43452	Sequence 43452, A
145	35	46.1	223	4	US-09-513-999C-4790	Sequence 4790, Ap	217	34	44.7	502	4	US-09-885-723-13	Sequence 13, Appl
146	35	46.1	223	4	US-09-513-999C-4793	Sequence 4793, Ap	218	34	44.7	502	4	US-09-885-723-13	Sequence 13, Appl
147	35	46.1	236	4	US-09-902-540-11067	Sequence 11067, A	219	34	44.7	502	4	US-09-252-991A-28225	Sequence 28225, A
148	35	46.1	263	4	US-09-902-540-12703	Sequence 12703, A	220	34	44.7	477	4	US-09-252-991A-28225	Sequence 28225, A
149	35	46.1	289	4	US-09-489-039A-12390	Sequence 12390, A	221	34	44.7	483	2	US-08-770-544-8	Sequence 8, Appli
150	35	46.1	344	4	US-09-270-767-58767	Sequence 58767, A	222	34	44.7	483	2	US-08-770-544-8	Sequence 8, Appli
151	35	46.1	399	4	US-09-252-991A-16821	Sequence 16821, A	223	34	44.7	483	2	US-08-770-544-8	Sequence 8, Appli
152	35	46.1	413	4	US-09-328-352-4779	Sequence 4779, Ap	224	34	44.7	483	2	US-08-770-544-8	Sequence 8, Appli
153	35	46.1	460	4	US-07-794-393-2	Sequence 39, Appl	225	34	44.7	483	2	US-08-770-544-8	Sequence 8, Appli
154	35	46.1	488	1	US-07-794-393-2	Sequence 39, Appl	226	34	44.7	483	2	US-08-770-544-8	Sequence 8, Appli
155	35	46.1	488	1	US-08-001-711-2	Sequence 2, Appli	227	34	44.7	494	4	US-09-579-259-8	Sequence 8, Appli
156	35	46.1	488	3	US-08-704-711A-22	Sequence 22, Appl	228	34	44.7	494	4	US-09-579-259-8	Sequence 8, Appli
157	35	46.1	488	3	US-09-521-220-22	Sequence 22, Appl	229	34	44.7	502	4	US-09-579-259-8	Sequence 8, Appli
158	35	46.1	488	3	US-09-391-104-31	Sequence 31, Appl	230	34	44.7	502	4	US-09-579-259-8	Sequence 8, Appli
159	35	46.1	489	3	US-08-448-489-11	Sequence 11, Appl	231	34	44.7	502	4	US-09-579-259-8	Sequence 8, Appli
160	35	46.1	489	4	US-09-689-730-11	Sequence 11, Appl	232	34	44.7	502	4	US-09-579-259-8	Sequence 8, Appli
161	35	46.1	490	4	US-09-902-540-10306	Sequence 10306, A	233	34	44.7	502	4	US-09-579-259-8	Sequence 8, Appli
162	35	46.1	491	4	US-09-902-540-10808	Sequence 10808, A	234	34	44.7	502	4	US-09-579-259-8	Sequence 8, Appli
163	35	46.1	497	4	US-08-956-171E-5234	Sequence 5234, Ap	235	34	44.7	502	4	US-09-579-259-8	Sequence 8, Appli
164	35	46.1	497	4	US-08-781-986A-5234	Sequence 5234, Ap	236	34	44.7	502	4	US-09-579-259-8	Sequence 8, Appli
165	35	46.1	500	4	US-09-252-991A-21214	Sequence 21214, A	237	34	44.7	502	4	US-09-579-259-8	Sequence 8, Appli
166	35	46.1	516	3	US-09-215-694-16	Sequence 16, Appl	238	34	44.7	502	4	US-09-579-259-8	Sequence 8, Appli
167	35	46.1	534	4	US-09-489-039A-10516	Sequence 10516, A	239	34	44.7	502	4	US-09-579-259-8	Sequence 8, Appli
168	35	46.1	546	4	US-09-902-540-15239	Sequence 15239, A	240	34	44.7	502	4	US-09-579-259-8	Sequence 8, Appli
169	35	46.1	592	4	US-09-813-453B-22	Sequence 22, Appl	241	34	44.7	502	4	US-09-579-259-8	Sequence 8, Appli
170	35	46.1	621	4	US-09-270-767-43417	Sequence 43417, A	242	34	44.7	502	4	US-09-579-259-8	Sequence 8, Appli
171	35	46.1	682	4	US-09-489-039A-7480	Sequence 7480, Ap	243	33	43.4	502	4	US-09-579-259-8	Sequence 8, Appli
172	35	46.1	760	4	US-09-252-991A-18711	Sequence 18711, A	244	33	43.4	502	4	US-09-579-259-8	Sequence 8, Appli
173	35	46.1	802	4	US-09-252-991A-23824	Sequence 23824, A	245	33	43.4	502	4	US-09-579-259-8	Sequence 8, Appli



247	33	43.4	67	4	US-09-543-681A-5000	Sequence 5000, Ap	320	33	43.4	398	1	US-08-476-000-4	Sequence 4, Appli
248	33	43.4	71	4	US-09-462-917A-120	Sequence 120, App	321	33	43.4	398	1	US-08-476-000-6	Sequence 6, Appli
249	33	43.4	72	4	US-09-540-236-2199	Sequence 2199, Ap	322	33	43.4	398	1	US-08-472-840-4	Sequence 4, Appli
250	33	43.4	76	4	US-09-328-352-2199	Sequence 4331, Ap	323	33	43.4	398	1	US-08-472-840-6	Sequence 6, Appli
251	33	43.4	76	4	US-09-328-352-2199	Sequence 4331, Ap	324	33	43.4	398	2	US-08-476-976-4	Sequence 4, Appli
252	33	43.4	93	3	US-08-981-397-59232	Sequence 59232, A	325	33	43.4	398	2	US-08-476-976-6	Sequence 6, Appli
253	33	43.4	93	3	US-09-908-322-44	Sequence 44, Appl	326	33	43.4	398	3	US-08-474-410-4	Sequence 4, Appli
254	33	43.4	112	4	US-09-543-681A-4261	Sequence 4261, Ap	327	33	43.4	398	3	US-08-474-410-6	Sequence 6, Appli
255	33	43.4	121	4	US-09-186-350A-11	Sequence 11, Appl	328	33	43.4	398	3	US-08-486-673B-4	Sequence 4, Appli
256	33	43.4	122	4	US-09-252-991A-30865	Sequence 30865, A	329	33	43.4	398	3	US-08-486-673B-6	Sequence 6, Appli
257	33	43.4	124	4	US-09-489-039A-11885	Sequence 11885, A	330	33	43.4	407	4	US-09-252-991A-24138	Sequence 24138, A
258	33	43.4	135	4	US-09-270-767-58995	Sequence 58995, A	331	33	43.4	408	4	US-09-949-016-9941	Sequence 9941, Ap
259	33	43.4	156	4	US-09-902-540-13491	Sequence 13491, A	332	33	43.4	410	3	US-09-205-283-2	Sequence 2, Appli
260	33	43.4	163	4	US-09-134-000C-4018	Sequence 4018, Ap	333	33	43.4	410	3	US-09-205-283-14	Sequence 14, Appl
261	33	43.4	169	4	US-09-489-039A-10447	Sequence 10447, A	334	33	43.4	421	3	US-08-484-993B-6	Sequence 6, Appli
262	33	43.4	179	4	US-09-134-000C-6423	Sequence 6423, Ap	335	33	43.4	421	2	US-08-484-158B-6	Sequence 6, Appli
263	33	43.4	191	4	US-09-270-767-36747	Sequence 36747, A	336	33	43.4	421	2	US-08-484-158B-8	Sequence 8, Appli
264	33	43.4	191	4	US-09-270-767-51964	Sequence 51964, A	337	33	43.4	421	2	US-08-480-150A-6	Sequence 6, Appli
265	33	43.4	202	4	US-09-252-991A-22744	Sequence 22744, A	338	33	43.4	421	3	US-08-458-731-6	Sequence 6, Appli
266	33	43.4	223	4	US-09-252-991A-20177	Sequence 20177, A	339	33	43.4	421	3	US-08-149-223A-6	Sequence 6, Appli
267	33	43.4	224	4	US-09-198-452A-583	Sequence 583, App	340	33	43.4	421	4	US-09-489-039A-7512	Sequence 7512, Ap
268	33	43.4	229	4	US-09-438-185A-648	Sequence 648, App	341	33	43.4	425	4	US-09-583-110-4415	Sequence 4415, Ap
269	33	43.4	234	3	US-09-040-483-5	Sequence 5, Appli	342	33	43.4	426	4	US-09-489-039A-11995	Sequence 11995, A
270	33	43.4	234	4	US-08-740-036-5	Sequence 42925, A	343	33	43.4	428	4	US-09-107-433-4773	Sequence 4773, Ap
271	33	43.4	256	4	US-09-270-767-42925	Sequence 42925, A	344	33	43.4	461	4	US-09-107-532A-5964	Sequence 5964, Ap
272	33	43.4	257	1	US-08-378-761A-13	Sequence 13, Appl	345	33	43.4	461	4	US-09-489-039A-12176	Sequence 12176, A
273	33	43.4	257	1	US-08-485-286-13	Sequence 13, Appl	346	33	43.4	464	1	US-07-688-352C-16	Sequence 16, Appl
274	33	43.4	258	3	US-09-097-767A-6	Sequence 6, Appli	347	33	43.4	464	2	US-08-474-379C-16	Sequence 16, Appl
275	33	43.4	270	3	US-09-323-872A-38	Sequence 38, Appl	348	33	43.4	464	3	US-09-146-249A-16	Sequence 16, Appl
276	33	43.4	292	4	US-09-538-092-608	Sequence 608, App	349	33	43.4	464	3	US-08-206-188B-16	Sequence 16, Appl
277	33	43.4	298	4	US-09-107-532A-4917	Sequence 4917, Ap	350	33	43.4	464	5	PCT-US91-02714-16	Sequence 16, Appl
278	33	43.4	305	4	US-09-252-991A-19225	Sequence 19225, A	351	33	43.4	468	4	US-09-603-787A-324	Sequence 324, App
279	33	43.4	308	4	US-09-804-778A-4	Sequence 4, Appli	352	33	43.4	477	4	US-09-252-991A-27656	Sequence 27656, A
280	33	43.4	313	4	US-09-248-796A-17621	Sequence 17621, A	353	33	43.4	494	3	US-09-205-283-12	Sequence 12, Appl
281	33	43.4	319	4	US-09-252-991A-29542	Sequence 29542, A	354	33	43.4	508	4	US-09-489-039A-7402	Sequence 7402, Ap
282	33	43.4	324	4	US-09-498-520A-2	Sequence 2, Appli	355	33	43.4	528	4	US-09-543-681A-6551	Sequence 6551, Ap
283	33	43.4	324	4	US-09-583-110-3913	Sequence 3913, Ap	356	33	43.4	537	4	US-09-489-039A-12290	Sequence 12290, A
284	33	43.4	324	4	US-09-968-129-2	Sequence 2, Appli	357	33	43.4	541	4	US-09-252-991A-27169	Sequence 27169, A
285	33	43.4	327	4	US-09-252-991A-19369	Sequence 19369, A	358	33	43.4	542	4	US-09-248-796A-15128	Sequence 15128, A
286	33	43.4	330	3	US-08-690-347-2	Sequence 2, Appli	359	33	43.4	552	4	US-09-949-016-11728	Sequence 11728, A
287	33	43.4	331	4	US-09-107-433-5083	Sequence 5083, Ap	360	33	43.4	585	4	US-09-252-991A-17061	Sequence 17061, A
288	33	43.4	336	4	US-09-603-208A-102	Sequence 102, App	361	33	43.4	585	4	US-09-248-796A-14925	Sequence 14925, A
289	33	43.4	345	4	US-09-902-540-12122	Sequence 12122, A	362	33	43.4	596	4	US-09-902-540-13547	Sequence 13547, A
290	33	43.4	360	4	US-09-252-991A-17420	Sequence 17420, A	363	33	43.4	603	4	US-09-252-991A-31905	Sequence 31905, A
291	33	43.4	365	3	US-08-155-005A-8	Sequence 8, Appli	364	33	43.4	604	1	US-08-224-657-84	Sequence 84, Appl
292	33	43.4	365	3	US-09-363-783-8	Sequence 8, Appli	365	33	43.4	604	1	US-08-224-657-95	Sequence 95, Appl
293	33	43.4	365	4	US-09-661-758A-8	Sequence 8, Appli	366	33	43.4	604	1	US-08-224-657-98	Sequence 98, Appl
294	33	43.4	368	2	US-08-824-878-3	Sequence 3, Appli	367	33	43.4	604	3	US-09-354-138-84	Sequence 84, Appl
295	33	43.4	368	2	US-09-353-688-3	Sequence 3, Appli	368	33	43.4	604	3	US-09-354-138-95	Sequence 95, Appl
296	33	43.4	371	4	US-09-540-236-3776	Sequence 3776, Ap	369	33	43.4	604	3	US-09-354-138-98	Sequence 98, Appl
297	33	43.4	373	1	US-08-118-270-24	Sequence 24, Appl	370	33	43.4	629	4	US-09-602-787A-624	Sequence 624, App
298	33	43.4	373	2	US-08-824-878-1	Sequence 1, Appli	371	33	43.4	667	4	US-09-747-259-16	Sequence 16, Appl
299	33	43.4	373	3	US-09-353-688-1	Sequence 1, Appli	372	33	43.4	667	4	US-09-816-744-16	Sequence 16, Appl
300	33	43.4	373	3	US-09-252-991A-20424	Sequence 20424, A	373	33	43.4	679	4	US-09-949-016-7268	Sequence 7268, Ap
301	33	43.4	373	5	PCT-US93-08528-24	Sequence 24, Appl	374	33	43.4	837	2	US-08-680-326-117	Sequence 117, App
302	33	43.4	388	3	US-08-155-005A-6	Sequence 6, Appli	375	33	43.4	899	4	US-09-902-540-13735	Sequence 13735, A
303	33	43.4	388	3	US-09-363-783-6	Sequence 6, Appli	376	33	43.4	955	2	US-08-500-857A-10	Sequence 10, Appl
304	33	43.4	388	4	US-09-661-758A-6	Sequence 6, Appli	377	33	43.4	956	4	US-09-949-016-8159	Sequence 8159, Ap
305	33	43.4	390	3	US-08-155-005A-4	Sequence 4, Appli	378	33	43.4	967	4	US-09-543-681A-6407	Sequence 6407, Ap
306	33	43.4	390	3	US-08-155-005A-17	Sequence 17, Appl	379	33	43.4	969	4	US-09-252-991A-26985	Sequence 26985, A
307	33	43.4	390	3	US-09-363-783-4	Sequence 4, Appli	380	33	43.4	1034	4	US-09-252-991A-28921	Sequence 28921, A
308	33	43.4	390	3	US-09-363-783-17	Sequence 17, Appl	381	33	43.4	1093	4	US-09-252-991A-21827	Sequence 21827, A
309	33	43.4	390	4	US-09-661-758A-4	Sequence 4, Appli	382	33	43.4	1150	3	US-09-238-303-9	Sequence 9, Appli
310	33	43.4	390	4	US-09-661-758A-17	Sequence 17, Appl	383	33	43.4	1150	4	US-09-946-239-9	Sequence 9, Appli
311	33	43.4	394	2	US-08-742-440A-8	Sequence 8, Appli	384	33	43.4	1233	4	US-09-134-000C-4971	Sequence 4971, Ap
312	33	43.4	397	1	US-08-476-000-63	Sequence 63, Appl	385	33	43.4	1246	4	US-09-919-497-85	Sequence 85, Appl
313	33	43.4	397	1	US-08-472-840-63	Sequence 63, Appl	386	33	43.4	1247	4	US-09-961-403-14	Sequence 14, Appl
314	33	43.4	397	2	US-08-476-976-63	Sequence 63, Appl	387	33	43.4	1338	4	US-09-631-603-2	Sequence 2, Appli
315	33	43.4	397	3	US-08-474-410-63	Sequence 63, Appl	388	33	43.4	1448	4	US-09-071-035-402	Sequence 402, App
316	33	43.4	397	3	US-08-486-673B-63	Sequence 63, Appl	389	33	43.4	1540	4	US-09-949-016-11382	Sequence 11382, A
317	33	43.4	397	4	US-09-602-787A-628	Sequence 628, App	390	33	43.4	1540	4	US-09-949-016-11383	Sequence 11383, A
318	33	43.4	398	1	US-08-097-938-6	Sequence 4, Appli	391	33	43.4	1719	2	US-08-459-568-4	Sequence 4, Appli
319	33	43.4	398	1	US-08-097-938-6	Sequence 6, Appli	392	33	43.4	1719	2	US-08-399-411-4	Sequence 4, Appli

393	33	43.4	1719	3	US-08-516-859A-4	Sequence 4, Appl1	466	32	42.1	181	4	US-09-328-352-4463	Sequence 4463, Ap
394	33	43.4	1719	3	US-09-586-472-4	Sequence 4, Appl1	467	32	42.1	184	4	US-09-583-110-4505	Sequence 4505, Ap
395	33	43.4	1719	4	US-09-528-706-4	Sequence 4, Appl1	468	32	42.1	188	4	US-09-270-767-3150	Sequence 3150, A
396	33	43.4	2216	4	US-09-502-540-12221	Sequence 12221, A	469	32	42.1	188	4	US-09-270-767-48367	Sequence 48367, A
397	33	43.4	4872	4	US-09-424-783-3	Sequence 3, Appl1	470	32	42.1	196	4	US-09-252-991A-19655	Sequence 19655, A
398	32.5	42.8	30	3	US-09-051-934-53	Sequence 53, Appl	471	32	42.1	205	2	US-08-775-009-37	Sequence 37, Appl
399	32.5	42.8	30	3	US-09-051-934-54	Sequence 54, Appl	472	32	42.1	205	4	US-09-949-016-11514	Sequence 11514, A
400	32.5	42.8	146	3	US-09-373-750-3	Sequence 3, Appl1	473	32	42.1	211	4	US-09-270-767-43785	Sequence 43785, A
401	32.5	42.8	254	4	US-09-252-991A-33117	Sequence 33117, A	474	32	42.1	214	4	US-09-252-991A-25513	Sequence 25513, A
402	32.5	42.8	494	4	US-09-252-991A-33227	Sequence 33227, A	475	32	42.1	217	4	US-09-270-767-34713	Sequence 34713, A
403	32.5	42.8	881	4	US-09-489-039A-13851	Sequence 13851, A	476	32	42.1	217	4	US-09-270-767-49930	Sequence 49930, A
404	32	42.1	16	4	US-09-556-877-138	Sequence 138, App	477	32	42.1	229	4	US-09-489-039A-14111	Sequence 14111, A
405	32	42.1	16	4	US-09-620-412C-138	Sequence 138, App	478	32	42.1	260	4	US-09-252-991A-26958	Sequence 26958, A
406	32	42.1	16	4	US-09-598-419-138	Sequence 138, App	479	32	42.1	272	3	US-09-461-697-186	Sequence 186, App
407	32	42.1	25	4	US-09-692-170C-25	Sequence 138, App	480	32	42.1	272	4	US-09-270-767-49903	Sequence 49903, A
408	32	42.1	25	4	US-10-405-231A-25	Sequence 25, Appl	481	32	42.1	283	4	US-09-145-828A-22	Sequence 22, Appl
409	32	42.1	25	4	US-10-238-607-25	Sequence 25, Appl	482	32	42.1	283	4	US-09-903-456-29	Sequence 29, Appl
410	32	42.1	25	4	US-09-984-365-25	Sequence 25, Appl	483	32	42.1	286	4	US-09-540-236-3041	Sequence 3041, Ap
411	32	42.1	30	4	US-09-471-276-1434	Sequence 1434, Ap	484	32	42.1	291	4	US-09-252-991A-28294	Sequence 28294, A
412	32	42.1	52	4	US-09-556-877-158	Sequence 158, App	485	32	42.1	297	4	US-09-489-039A-7389	Sequence 7389, Ap
413	32	42.1	52	4	US-09-620-412C-158	Sequence 158, App	486	32	42.1	298	4	US-09-556-877-17	Sequence 17, Appl
414	32	42.1	52	4	US-09-598-419-158	Sequence 158, App	487	32	42.1	298	4	US-09-556-877-121	Sequence 121, App
415	32	42.1	53	4	US-09-556-877-32	Sequence 32, Appl	488	32	42.1	298	4	US-09-556-877-123	Sequence 123, App
416	32	42.1	53	4	US-09-556-877-34	Sequence 34, Appl	489	32	42.1	298	4	US-09-556-877-125	Sequence 125, App
417	32	42.1	53	4	US-09-288-594A-32	Sequence 34, Appl	490	32	42.1	298	4	US-09-556-877-127	Sequence 127, App
418	32	42.1	53	4	US-09-288-594A-34	Sequence 34, Appl	491	32	42.1	298	4	US-09-556-877-129	Sequence 129, App
419	32	42.1	53	4	US-09-620-412C-32	Sequence 32, Appl	492	32	42.1	298	4	US-09-556-877-131	Sequence 131, App
420	32	42.1	53	4	US-09-620-412C-34	Sequence 34, Appl	493	32	42.1	298	4	US-09-556-877-133	Sequence 133, App
421	32	42.1	53	4	US-09-410-568-32	Sequence 32, Appl	494	32	42.1	298	4	US-09-556-877-135	Sequence 135, App
422	32	42.1	53	4	US-09-410-568-34	Sequence 34, Appl	495	32	42.1	298	4	US-09-556-877-264	Sequence 264, App
423	32	42.1	53	4	US-09-598-419-32	Sequence 32, Appl	496	32	42.1	298	4	US-09-556-877-266	Sequence 266, App
424	32	42.1	53	4	US-09-598-419-34	Sequence 34, Appl	497	32	42.1	298	4	US-09-288-594A-17	Sequence 17, Appl
425	32	42.1	58	4	US-09-621-976-5217	Sequence 5217, Ap	498	32	42.1	298	4	US-09-620-412C-17	Sequence 17, Appl
426	32	42.1	63	4	US-09-621-976-4480	Sequence 4480, Ap	499	32	42.1	298	4	US-09-620-412C-121	Sequence 121, App
427	32	42.1	70	4	US-09-732-210-1543	Sequence 1543, Ap	500	32	42.1	298	4	US-09-620-412C-123	Sequence 123, App
428	32	42.1	72	1	US-08-480-443-11	Sequence 11, Appl	501	32	42.1	298	4	US-09-620-412C-125	Sequence 125, App
429	32	42.1	72	1	US-08-457-459-11	Sequence 11, Appl	502	32	42.1	298	4	US-09-620-412C-127	Sequence 127, App
430	32	42.1	72	1	US-08-555-678-11	Sequence 11, Appl	503	32	42.1	298	4	US-09-620-412C-129	Sequence 129, App
431	32	42.1	72	4	US-09-270-767-37426	Sequence 37426, A	504	32	42.1	298	4	US-09-620-412C-131	Sequence 131, App
432	32	42.1	72	4	US-09-270-767-52643	Sequence 52643, A	505	32	42.1	298	4	US-09-620-412C-133	Sequence 133, App
433	32	42.1	72	5	PCT-US95-02275-11	Sequence 11, Appl	506	32	42.1	298	4	US-09-620-412C-135	Sequence 135, App
434	32	42.1	79	4	US-09-621-976-4478	Sequence 4478, Ap	507	32	42.1	298	4	US-09-620-412C-264	Sequence 264, App
435	32	42.1	79	4	US-09-248-796A-27215	Sequence 27215, A	508	32	42.1	298	4	US-09-620-412C-266	Sequence 266, App
436	32	42.1	80	4	US-09-621-976-4479	Sequence 4479, Ap	509	32	42.1	298	4	US-09-410-568-17	Sequence 17, Appl
437	32	42.1	85	4	US-09-513-990C-4697	Sequence 4697, Ap	510	32	42.1	298	4	US-09-598-419-17	Sequence 17, Appl
438	32	42.1	85	4	US-09-471-276-880	Sequence 880, App	511	32	42.1	298	4	US-09-598-419-121	Sequence 121, App
439	32	42.1	99	4	US-09-513-999C-7494	Sequence 7494, Ap	512	32	42.1	298	4	US-09-598-419-123	Sequence 123, App
440	32	42.1	101	4	US-09-710-379-280	Sequence 280, App	513	32	42.1	298	4	US-09-598-419-125	Sequence 125, App
441	32	42.1	105	4	US-09-489-039A-11498	Sequence 11498, A	514	32	42.1	298	4	US-09-598-419-127	Sequence 127, App
442	32	42.1	115	4	US-09-270-767-32320	Sequence 32320, A	515	32	42.1	298	4	US-09-598-419-129	Sequence 129, App
443	32	42.1	117	4	US-09-270-767-61108	Sequence 61108, A	516	32	42.1	298	4	US-09-598-419-131	Sequence 131, App
444	32	42.1	117	4	US-09-949-016-11125	Sequence 11125, A	517	32	42.1	298	4	US-09-598-419-133	Sequence 133, App
445	32	42.1	121	4	US-09-621-976-6925	Sequence 6925, Ap	518	32	42.1	298	4	US-09-598-419-135	Sequence 135, App
446	32	42.1	126	4	US-09-902-540-11925	Sequence 11925, A	519	32	42.1	298	4	US-09-598-419-264	Sequence 264, App
447	32	42.1	130	4	US-09-270-767-60369	Sequence 60369, A	520	32	42.1	298	4	US-09-598-419-266	Sequence 266, App
448	32	42.1	130	4	US-09-949-016-8171	Sequence 8171, Ap	521	32	42.1	303	4	US-09-543-681A-5539	Sequence 5539, Ap
449	32	42.1	132	4	US-09-252-991A-29618	Sequence 29618, A	522	32	42.1	307	4	US-09-489-039A-8974	Sequence 8974, Ap
450	32	42.1	135	4	US-09-270-767-37619	Sequence 37619, A	523	32	42.1	308	4	US-09-804-778A-2	Sequence 2, Appl1
451	32	42.1	135	4	US-09-270-767-52836	Sequence 52836, A	524	32	42.1	308	4	US-09-907-794A-153	Sequence 153, App
452	32	42.1	136	4	US-09-252-991A-17270	Sequence 17270, A	525	32	42.1	310	4	US-09-905-125A-153	Sequence 153, App
453	32	42.1	137	4	US-09-583-110-4762	Sequence 4762, Ap	526	32	42.1	310	4	US-09-902-775A-153	Sequence 153, App
454	32	42.1	141	4	US-09-252-991A-27362	Sequence 27362, A	527	32	42.1	310	4	US-09-906-706-153	Sequence 153, App
455	32	42.1	143	4	US-08-529-055-38	Sequence 38, Appl	528	32	42.1	310	4	US-09-903-603A-133	Sequence 133, App
456	32	42.1	143	4	US-09-270-767-56776	Sequence 56776, A	529	32	42.1	310	4	US-09-904-920A-153	Sequence 153, App
457	32	42.1	143	4	US-09-107-433-2613	Sequence 2613, Ap	530	32	42.1	310	4	US-09-909-064-153	Sequence 153, App
458	32	42.1	151	4	US-09-252-991A-18738	Sequence 18738, A	531	32	42.1	310	4	US-09-905-381A-153	Sequence 153, App
459	32	42.1	152	4	US-09-311-784A-20	Sequence 20, Appl	532	32	42.1	310	4	US-09-906-618-153	Sequence 153, App
460	32	42.1	156	4	US-09-252-991A-32181	Sequence 32181, A	533	32	42.1	313	4	US-09-949-016-7795	Sequence 7795, Ap
461	32	42.1	158	4	US-09-252-991A-28052	Sequence 28052, A	534	32	42.1	313	4	US-08-680-506-3	Sequence 3, Appl1
462	32	42.1	159	4	US-09-543-681A-6783	Sequence 6783, Ap	535	32	42.1	318	3	US-09-252-991A-21237	Sequence 21237, A
463	32	42.1	161	4	US-09-270-767-45593	Sequence 45593, A	536	32	42.1	321	4	US-09-311-021-74	Sequence 74, Appl
464	32	42.1	177	4	US-09-270-767-35220	Sequence 35220, A	537	32	42.1	325	4	US-09-806-536A-10	Sequence 10, Appl
465	32	42.1	177	4	US-09-270-767-50437	Sequence 50437, A	538	32	42.1	325	4		

539	32	42.1	327	4	US-09-252-991A-30065	Sequence 30065, A	612	32	42.1	479	3	US-08-767-993-7	Sequence 7, Appli
540	32	42.1	330	2	US-08-827-615-2	Sequence 2, Appli	613	32	42.1	479	4	US-09-107-532A-7052	Sequence 7052, Ap
541	32	42.1	330	3	US-08-921-209-4	Sequence 4, Appli	614	32	42.1	479	4	US-09-134-000C-4707	Sequence 4707, Ap
542	32	42.1	330	3	US-09-411-763-4	Sequence 4, Appli	615	32	42.1	483	4	US-09-489-039A-13018	Sequence 13018, A
543	32	42.1	337	4	US-09-543-681A-6463	Sequence 6463, Ap	616	32	42.1	504	4	US-09-602-787A-500	Sequence 500, App
544	32	42.1	342	1	US-08-118-270-9	Sequence 9, Appli	617	32	42.1	507	1	US-08-484-493-12	Sequence 12, Appl
545	32	42.1	342	5	PCT-US93-08528-9	Sequence 9, Appli	618	32	42.1	507	1	US-08-484-493-12	Sequence 12, Appl
546	32	42.1	347	4	US-09-270-767-32014	Sequence 32014, A	619	32	42.1	507	2	US-08-345-213-12	Sequence 12, Appl
547	32	42.1	347	4	US-09-270-767-47231	Sequence 47231, A	620	32	42.1	507	3	US-09-249-003-12	Sequence 12, Appl
548	32	42.1	350	4	US-09-149-476-452	Sequence 452, App	621	32	42.1	507	3	US-09-685-844-12	Sequence 12, Appl
549	32	42.1	351	4	US-09-489-039A-12237	Sequence 12237, A	622	32	42.1	509	4	US-10-142-231-67	Sequence 67, Appl
550	32	42.1	355	1	US-08-118-270-11	Sequence 11, Appl	623	32	42.1	511	4	US-09-252-991A-18418	Sequence 18418, A
551	32	42.1	355	5	PCT-US93-08528-11	Sequence 11, Appl	624	32	42.1	530	4	US-09-252-991A-21965	Sequence 21965, A
552	32	42.1	365	2	US-08-855-518-5	Sequence 5, Appli	625	32	42.1	531	4	US-09-248-796A-18979	Sequence 18979, A
553	32	42.1	365	4	US-09-252-991A-21955	Sequence 21955, A	626	32	42.1	538	4	US-09-107-532A-5563	Sequence 5563, Ap
554	32	42.1	366	2	US-08-855-518-4	Sequence 4, Appli	627	32	42.1	538	4	US-09-489-039A-8363	Sequence 8363, Ap
555	32	42.1	376	4	US-09-270-767-41550	Sequence 41550, A	628	32	42.1	563	1	US-08-442-884-1	Sequence 1, Appli
556	32	42.1	379	4	US-09-252-991A-21117	Sequence 21117, A	629	32	42.1	564	4	US-09-252-991A-25356	Sequence 25356, A
557	32	42.1	384	4	US-09-949-016-6494	Sequence 6494, Ap	630	32	42.1	565	4	US-08-937-067-8	Sequence 8, Appli
558	32	42.1	388	1	US-08-087-772A-2	Sequence 2, Appli	631	32	42.1	568	4	US-09-328-352-5502	Sequence 5502, Ap
559	32	42.1	388	4	US-09-583-110-3393	Sequence 3393, Ap	632	32	42.1	568	4	US-09-690-942-6	Sequence 6, Appli
560	32	42.1	389	4	US-09-489-039A-12200	Sequence 12200, A	633	32	42.1	568	4	US-09-690-942-8	Sequence 8, Appli
561	32	42.1	390	4	US-09-107-532A-5569	Sequence 5569, Ap	634	32	42.1	570	4	US-09-690-942-14	Sequence 14, Appl
562	32	42.1	392	4	US-09-252-991A-21979	Sequence 21979, A	635	32	42.1	581	4	US-09-252-991A-30512	Sequence 30512, A
563	32	42.1	396	4	US-09-107-433-4213	Sequence 4213, Ap	636	32	42.1	585	4	US-09-489-039A-14148	Sequence 14148, A
564	32	42.1	397	4	US-09-543-681A-4786	Sequence 4786, Ap	637	32	42.1	588	4	US-09-690-942-13	Sequence 13, Appl
565	32	42.1	400	1	US-07-916-901-6	Sequence 6, Appli	638	32	42.1	590	4	US-09-538-092-967	Sequence 967, App
566	32	42.1	400	1	US-07-783-602C-1	Sequence 1, Appli	639	32	42.1	605	2	US-08-472-666-1	Sequence 1, Appli
567	32	42.1	400	1	US-08-351-473B-4	Sequence 4, Appli	640	32	42.1	605	5	PCT-US96-07615-1	Sequence 1, Appli
568	32	42.1	400	1	US-08-351-473B-5	Sequence 5, Appli	641	32	42.1	616	4	US-09-543-681A-4421	Sequence 4421, Ap
569	32	42.1	400	3	US-08-450-962-4	Sequence 4, Appli	642	32	42.1	632	4	US-09-661-711A-5	Sequence 5, Appli
570	32	42.1	400	3	US-08-450-962-6	Sequence 6, Appli	643	32	42.1	632	4	US-09-538-092-159	Sequence 159, App
571	32	42.1	400	4	US-08-848-631-4	Sequence 4, Appli	644	32	42.1	634	3	US-09-357-251-10	Sequence 10, Appl
572	32	42.1	400	4	US-08-848-631-6	Sequence 6, Appli	645	32	42.1	635	4	US-09-538-092-594	Sequence 594, App
573	32	42.1	401	4	US-09-489-039A-10007	Sequence 10007, A	646	32	42.1	659	4	US-09-543-681A-4748	Sequence 4748, Ap
574	32	42.1	405	1	US-08-351-473B-2	Sequence 2, Appli	647	32	42.1	676	4	US-09-270-767-44900	Sequence 44900, A
575	32	42.1	414	4	US-09-583-110-4310	Sequence 4310, Ap	648	32	42.1	681	4	US-09-270-767-44219	Sequence 44219, A
576	32	42.1	415	4	US-09-543-681A-6292	Sequence 6292, Ap	649	32	42.1	702	3	US-09-232-200-102	Sequence 102, App
577	32	42.1	415	4	US-09-107-433-3943	Sequence 3943, Ap	650	32	42.1	702	3	US-09-232-201-102	Sequence 102, App
578	32	42.1	420	4	US-09-252-991A-25846	Sequence 25846, A	651	32	42.1	702	3	US-09-232-201-102	Sequence 102, App
579	32	42.1	423	4	US-09-328-352-6273	Sequence 6273, Ap	652	32	42.1	716	4	US-09-489-039A-10841	Sequence 10841, A
580	32	42.1	426	4	US-09-252-991A-24450	Sequence 24450, A	653	32	42.1	723	4	US-09-949-016-9810	Sequence 9810, Ap
581	32	42.1	426	4	US-09-602-787A-424	Sequence 424, App	654	32	42.1	729	4	US-09-252-991A-16877	Sequence 16877, A
582	32	42.1	427	4	US-09-543-681A-6225	Sequence 6225, Ap	655	32	42.1	743	4	US-09-252-991A-16877	Sequence 16877, A
583	32	42.1	430	3	US-09-134-001C-2981	Sequence 2981, Ap	656	32	42.1	755	4	US-09-248-796A-15109	Sequence 15109, A
584	32	42.1	431	2	US-08-928-613-2	Sequence 2, Appli	657	32	42.1	789	4	US-09-134-000C-4939	Sequence 4939, Ap
585	32	42.1	431	3	US-09-193-524-2	Sequence 2, Appli	658	32	42.1	834	4	US-09-252-991A-17616	Sequence 17616, A
586	32	42.1	433	2	US-09-139-424-4	Sequence 4, Appli	659	32	42.1	868	1	US-08-374-834-1	Sequence 1, Appli
587	32	42.1	433	3	US-08-706-216-4	Sequence 4, Appli	660	32	42.1	868	2	US-08-644-271-1	Sequence 1, Appli
588	32	42.1	433	3	US-09-650-284B-4	Sequence 2, Appli	661	32	42.1	868	1	US-09-077-955-1	Sequence 1, Appli
589	32	42.1	436	3	US-08-660-347-2	Sequence 2, Appli	662	32	42.1	869	1	US-08-374-834-16	Sequence 16, Appl
590	32	42.1	439	3	US-08-921-209-2	Sequence 2, Appli	663	32	42.1	869	2	US-08-644-271-29	Sequence 29, Appl
591	32	42.1	439	3	US-09-411-763-2	Sequence 2, Appli	664	32	42.1	869	2	US-09-077-955-33	Sequence 33, Appl
592	32	42.1	440	3	US-09-631-603-22	Sequence 22, Appl	665	32	42.1	869	4	US-09-715-249-8	Sequence 8, Appli
593	32	42.1	440	4	US-09-826-509-567	Sequence 567, App	666	32	42.1	885	4	US-09-252-991A-26129	Sequence 26129, A
594	32	42.1	442	4	US-09-949-016-8625	Sequence 8625, Ap	667	32	42.1	907	4	US-09-198-452A-306	Sequence 306, App
595	32	42.1	445	4	US-09-583-110-4414	Sequence 4414, Ap	668	32	42.1	928	4	US-09-438-185A-295	Sequence 295, App
596	32	42.1	448	4	US-09-107-433-4864	Sequence 4864, Ap	669	32	42.1	1047	4	US-09-523-239D-99	Sequence 99, Appl
597	32	42.1	449	3	US-08-680-506-7	Sequence 7, Appli	670	32	42.1	1063	1	US-08-093-453B-3	Sequence 3, Appli
598	32	42.1	460	2	US-09-677-049-10	Sequence 10, Appl	671	32	42.1	1063	1	US-08-127-499A-8	Sequence 8, Appli
599	32	42.1	460	4	US-09-826-509-513	Sequence 513, App	672	32	42.1	1063	1	US-08-482-847-8	Sequence 8, Appli
600	32	42.1	461	1	US-08-196-989B-13	Sequence 13, Appl	673	32	42.1	1164	3	US-09-457-708-2	Sequence 2, Appli
601	32	42.1	461	2	US-08-760-936-13	Sequence 13, Appl	674	32	42.1	1164	4	US-09-950-046A-2	Sequence 2, Appli
602	32	42.1	461	4	US-09-225-024-13	Sequence 13, Appl	675	32	42.1	1164	4	US-09-976-594-989	Sequence 989, App
603	32	42.1	461	4	US-09-902-540-15885	Sequence 15885, A	676	32	42.1	1203	3	US-09-207-857-2	Sequence 2, Appli
604	32	42.1	463	4	US-09-252-991A-25488	Sequence 25488, A	677	32	42.1	1203	3	US-09-293-505-2	Sequence 2, Appli
605	32	42.1	470	4	US-09-543-681A-5952	Sequence 5952, Ap	678	32	42.1	1203	4	US-09-902-280A-2	Sequence 2, Appli
606	32	42.1	472	2	US-08-459-346-17	Sequence 17, Appl	679	32	42.1	1203	4	US-09-060-939A-2	Sequence 2, Appli
607	32	42.1	472	3	US-08-889-415-17	Sequence 17, Appl	680	32	42.1	1296	4	US-08-857-636-60	Sequence 60, Appl
608	32	42.1	472	3	US-08-402-542-17	Sequence 17, Appl	681	32	42.1	1434	2	US-08-540-406-10	Sequence 10, Appl
609	32	42.1	472	5	PCT-US93-07189-17	Sequence 17, Appl	682	32	42.1	1434	3	US-08-656-055-10	Sequence 10, Appl
610	32	42.1	478	4	US-09-489-039A-7205	Sequence 7205, Ap	683	32	42.1	1434	3	US-08-954-668-10	Sequence 10, Appl
611	32	42.1	479	1	US-08-313-553-7	Sequence 7, Appli	684	32	42.1	1434	4	US-08-918-658-10	Sequence 10, Appl

685	32	42.1	1434	4	US-09-724-631-10	Sequence 10, Appl
686	32	42.1	1434	4	US-08-354-701A-10	Sequence 10, Appl
687	32	42.1	1434	5	PCT-US95-13233-10	Sequence 10, Appl
688	32	42.1	1447	2	US-08-540-406-19	Sequence 19, Appl
689	32	42.1	1447	3	US-08-656-055-19	Sequence 19, Appl
690	32	42.1	1447	3	US-08-954-668-19	Sequence 19, Appl
691	32	42.1	1447	3	US-09-268-140-5	Sequence 5, Appl
692	32	42.1	1447	4	US-08-918-658-19	Sequence 19, Appl
693	32	42.1	1447	4	US-09-724-631-19	Sequence 19, Appl
694	32	42.1	1447	4	US-08-954-701A-19	Sequence 19, Appl
695	32	42.1	1447	5	PCT-US95-13233-19	Sequence 19, Appl
696	32	42.1	1493	3	US-09-376-330-16	Sequence 16, Appl
697	32	42.1	1969	4	US-09-418-710-72	Sequence 72, Appl
698	32	42.1	1969	4	US-09-839-479-71	Sequence 71, Appl
699	32	42.1	1972	4	US-09-418-710-21	Sequence 21, Appl
700	32	42.1	1972	4	US-09-839-479-21	Sequence 21, Appl
701	32	42.1	2628	3	US-09-413-814-11	Sequence 11, Appl
702	32	42.1	2890	3	US-09-413-814-67	Sequence 67, Appl
703	32	42.1	3798	3	US-09-335-409-6	Sequence 6, Appl
704	32	42.1	3798	3	US-09-568-102-6	Sequence 6, Appl
705	32	42.1	3798	3	US-09-567-969-6	Sequence 6, Appl
706	32	42.1	3798	3	US-09-568-480-6	Sequence 6, Appl
707	32	42.1	3798	3	US-09-568-486-6	Sequence 6, Appl
708	32	42.1	3798	3	US-09-568-472-6	Sequence 6, Appl
709	32	42.1	3798	3	US-09-567-899-6	Sequence 6, Appl
710	32	42.1	4866	4	US-09-424-783-2	Sequence 2, Appl
711	32	42.1	8991	4	US-08-714-741-32	Sequence 32, Appl
712	31.5	41.4	122	4	US-09-902-540-16535	Sequence 16535, A
713	31.5	41.4	144	4	US-09-749-340-2	Sequence 2, Appl
714	31.5	41.4	144	4	US-09-328-352-6154	Sequence 6154, Ap
715	31.5	41.4	378	4	US-09-489-039A-13960	Sequence 13960, A
716	31.5	41.4	659	4	US-09-252-991A-23467	Sequence 23467, A
717	31.5	41.4	4472	2	US-08-804-227C-2	Sequence 2, Appl
718	31	40.8	9	3	US-09-186-170-7	Sequence 7, Appl
719	31	40.8	9	4	US-09-562-868-7	Sequence 7, Appl
720	31	40.8	9	4	US-10-116-288A-7	Sequence 7, Appl
721	31	40.8	9	4	US-09-671-089-21	Sequence 21, Appl
722	31	40.8	11	4	US-09-671-089-12	Sequence 12, Appl
723	31	40.8	25	3	US-08-641-873-15	Sequence 15, Appl
724	31	40.8	27	4	US-09-936-885A-13	Sequence 13, Appl
725	31	40.8	28	4	US-09-270-767-62415	Sequence 62415, A
726	31	40.8	47	4	US-09-675-305-2	Sequence 2, Appl
727	31	40.8	47	4	US-10-200-344-2	Sequence 2, Appl
728	31	40.8	60	4	US-09-248-796A-21814	Sequence 21814, A
729	31	40.8	62	4	US-09-248-796A-24308	Sequence 24308, A
730	31	40.8	62	6	5496550-8	Patent No. 5496550
731	31	40.8	62	6	5496550-8	Patent No. 5496550
732	31	40.8	66	4	US-09-489-039A-10473	Sequence 10473, A
733	31	40.8	69	4	US-09-149-476-601	Sequence 601, App
734	31	40.8	70	4	US-09-149-476-465	Sequence 465, App
735	31	40.8	72	4	US-09-621-976-3951	Sequence 3951, Ap
736	31	40.8	75	4	US-09-513-999C-5662	Sequence 5662, Ap
737	31	40.8	76	4	US-09-248-796A-21942	Sequence 21942, A
738	31	40.8	76	4	US-09-248-796A-25392	Sequence 25392, A
739	31	40.8	77	4	US-09-248-796A-22112	Sequence 22112, A
740	31	40.8	82	4	US-09-270-767-44099	Sequence 44099, A
741	31	40.8	88	4	US-09-675-305-4	Sequence 4, Appl
742	31	40.8	88	4	US-10-200-344-4	Sequence 4, Appl
743	31	40.8	91	4	US-09-198-452A-465	Sequence 465, App
744	31	40.8	92	4	US-09-675-305-8	Sequence 8, Appl
745	31	40.8	92	4	US-10-200-344-8	Sequence 8, Appl
746	31	40.8	98	4	US-09-248-796A-18737	Sequence 18737, A
747	31	40.8	100	4	US-09-583-110-4764	Sequence 4764, Ap
748	31	40.8	101	4	US-09-107-433-3624	Sequence 3624, Ap
749	31	40.8	102	4	US-09-248-796A-14759	Sequence 14759, A
750	31	40.8	104	4	US-09-270-767-60187	Sequence 60187, A

ALIGNMENTS

; Sequence 24, Application US/09671089  
; Patent No. 6780846  
; GENERAL INFORMATION:  
; APPLICANT: O'Mahony, Daniel J.  
; APPLICANT: Lambkin, Inelda J.  
; TITLE OF INVENTION: MEMBRANE TRANSLOCATING PEPTIDE DRUG DELIVERY SYSTEM  
; FILE REFERENCE: E1067/20018  
; CURRENT APPLICATION NUMBER: US/09/671,089  
; CURRENT FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: 60/156,246  
; PRIOR FILING DATE: 1999-09-27  
; NUMBER OF SEQ ID NOS: 59  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 24  
; LENGTH: 17  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: membrane translocating peptide  
US-09-671-089-24

Query Match 100.0%; Score 76; DB 4; Length 17;  
Best Local Similarity 100.0%; Pred. No. 8.9e-06;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KKKKRAAAVLLPVLLA 17  
Db 1 KKKKRAAAVLLPVLLA 17

RESULT 2

US-08-928-958-24  
; Sequence 24, Application US/08928958  
; Patent No. 5877282  
; GENERAL INFORMATION:

; APPLICANT: NADLER, STEVEN G.  
; APPLICANT: CLEAVELAND, JEFFREY S.  
; APPLICANT: BLAKE, JAMES  
; APPLICANT: HAFAR, ONAR K.

; TITLE OF INVENTION: PEPTIDE INHIBITORS OF NUCLEAR PROTEIN  
; TITLE OF INVENTION: TRANSLOCATION HAVING NUCLEAR LOCALIZATION SEQUENCES AND  
; TITLE OF INVENTION: METHODS OF USE THEREOF  
; NUMBER OF SEQUENCES: 24  
; CORRESPONDENCE ADDRESS:

; ADDRESSEE: ROBINS & ASSOCIATES  
; STREET: 90 MIDDLEFIELD ROAD, SUITE 200  
; CITY: MENLO PARK  
; STATE: CA  
; COUNTRY: USA

; ZIP: 94025

; COMPUTER READABLE FORM: disk

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/928,958

; FILING DATE: 12-SEP-1997

; CLASSIFICATION: 514

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 60/026978

; FILING DATE: 20-SEP-1996

; ATTORNEY/AGENT INFORMATION:

; NAME: ROBINS, ROBERTA L.

; REGISTRATION NUMBER: 33,208

; REFERENCE/DOCKET NUMBER: 5998-0019

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (650) 325-7812

; TELEFAX: (650) 325-7823

; INFORMATION FOR SEQ ID NO: 24:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 29 amino acids

; TYPE: amino acid

```
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-928-958-24

Query Match      70.4%; Score 53.5; DB 2; Length 29;
Best Local Similarity 77.8%; Pred. No. 0.055;
Matches 14; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

Qy      1 KKKRKAATAVLLP-VLLA 17
        ||||| || |||||
Db      2 KKKRKAATAVLLP-VLLA 19

RESULT 3
US-09-072-429-24
; Sequence 24, Application US/09072429
; Patent No. 5962415
; GENERAL INFORMATION:
; APPLICANT: Nadler, Steven G.
; TITLE OF INVENTION: COMPOSITIONS COMPRISING A PEPTIDE
; TITLE OF INVENTION: INHIBITOR OF NUCLEAR PROTEIN TRANSLOCATION AND AN
; TITLE OF INVENTION: IMMUNOSUPPRESSANT AND METHODS OF USE THEREOF
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESS: Bristol-Myers Squibb Company
; STREET: P.O. Box 4000
; CITY: Princeton
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 08543-4000
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/072,429
; FILING DATE: 04-MAY-1998
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Klein, Christopher A.
; REGISTRATION NUMBER: 34,363
; REFERENCE/DOCKET NUMBER: ON0141b
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (609) 252-3714
; TELEFAX: (609) 252-4526
; INFORMATION FOR SEQ ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 29 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-072-429-24

Query Match      70.4%; Score 53.5; DB 2; Length 29;
Best Local Similarity 77.8%; Pred. No. 0.055;
Matches 14; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

Qy      1 KKKRKAATAVLLP-VLLA 17
        ||||| || |||||
Db      2 KKKRKAATAVLLP-VLLA 19

RESULT 4
US-08-928-958-1
; Sequence 1, Application US/08928958
; Patent No. 587282
; GENERAL INFORMATION:
; APPLICANT: NADLER, STEVEN G.
; APPLICANT: CLEVELAND, JEFFREY S.
; APPLICANT: BLAKE, JAMES
```

---

```
; APPLICANT: HAFAR, OMAR K.
; TITLE OF INVENTION: PEPTIDE INHIBITORS OF NUCLEAR PROTEIN
; TITLE OF INVENTION: TRANSLOCATION HAVING NUCLEAR LOCALIZATION SEQUENCES AND
; TITLE OF INVENTION: METHODS OF USE THEREOF
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESS: ROBINS & ASSOCIATES
; STREET: 90 MIDDLEFIELD ROAD, SUITE 200
; CITY: MENLO PARK
; STATE: CA
; COUNTRY: USA
; ZIP: 94025
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/928,958
; FILING DATE: 12-SEP-1997
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/026978
; FILING DATE: 20-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: ROBINS, ROBERTA L.
; REGISTRATION NUMBER: 33,208
; REFERENCE/DOCKET NUMBER: 5998-0019
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 325-7812
; TELEFAX: (650) 325-7823
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 30 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-928-958-1

Query Match      70.4%; Score 53.5; DB 2; Length 30;
Best Local Similarity 77.8%; Pred. No. 0.057;
Matches 14; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

Qy      1 KKKRKAATAVLLP-VLLA 17
        ||||| || |||||
Db      2 KKKRKAATAVLLP-VLLA 19

RESULT 5
US-09-072-429-1
; Sequence 1, Application US/09072429
; Patent No. 5962415
; GENERAL INFORMATION:
; APPLICANT: Nadler, Steven G.
; TITLE OF INVENTION: COMPOSITIONS COMPRISING A PEPTIDE
; TITLE OF INVENTION: INHIBITOR OF NUCLEAR PROTEIN TRANSLOCATION AND AN
; TITLE OF INVENTION: IMMUNOSUPPRESSANT AND METHODS OF USE THEREOF
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESS: Bristol-Myers Squibb Company
; STREET: P.O. Box 4000
; CITY: Princeton
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 08543-4000
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/072,429
```

```
; FILING DATE: 04-MAY-1998
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Klein, Christopher A.
; REGISTRATION NUMBER: 34,363
; REFERENCE/DOCKET NUMBER: ON0141b
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (609) 252-3714
; TELEFAX: (609) 252-4526
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 30 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-072-429-1
```

```
Query Match 70.4%; Score 53.5; DB 2; Length 30;
Best Local Similarity 77.8%; Pred. No. 0.057;
Matches 14; Conservative 0; Mismatches 3; Indels 1; Gaps 1;
```

```
Qy 1 KKKKRAAAVLLP-VLLA 17
Db 2 KKKRKAVALLPVALLA 19
```

```
RESULT 6
US-09-671-089-3
; Sequence 3, Application US/09671089
; Patent No. 6780846
; GENERAL INFORMATION:
; APPLICANT: O'Mahony, Daniel J.
; TITLE OF INVENTION: MEMBRANE TRANSLOCATING PEPTIDE DRUG DELIVERY SYSTEM
; FILE REFERENCE: E1067/20018
; CURRENT APPLICATION NUMBER: US/09/671,089
; CURRENT FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/156,246
; PRIOR FILING DATE: 1999-09-27
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 3
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: membrane translocating peptide
US-09-671-089-3
```

```
Query Match 69.7%; Score 53; DB 4; Length 16;
Best Local Similarity 85.7%; Pred. No. 0.036;
Matches 12; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

```
Qy 4 KRAAAVLLP-VLLA 17
Db 1 KKKAAVLLP-VLLA 14
```

```
RESULT 7
US-09-671-089-48
; Sequence 48, Application US/09671089
; Patent No. 6780846
; GENERAL INFORMATION:
; APPLICANT: O'Mahony, Daniel J.
; TITLE OF INVENTION: MEMBRANE TRANSLOCATING PEPTIDE DRUG DELIVERY SYSTEM
; FILE REFERENCE: E1067/20018
; CURRENT APPLICATION NUMBER: US/09/671,089
; CURRENT FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/156,246
; PRIOR FILING DATE: 1999-09-27
; NUMBER OF SEQ ID NOS: 59
```

```
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 48
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: dansylated membrane translocating peptide
; NAME/KEY: MOD_RES
; LOCATION: (1)..(1)
; OTHER INFORMATION: dansylated
US-09-671-089-48
```

```
Query Match 69.7%; Score 53; DB 4; Length 16;
Best Local Similarity 85.7%; Pred. No. 0.036;
Matches 12; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 4 KRAAAVLLP-VLLA 17
Db 1 KKKAAVLLP-VLLA 14
```

```
RESULT 8
US-09-671-089-2
; Sequence 2, Application US/09671089
; Patent No. 6780846
; GENERAL INFORMATION:
; APPLICANT: O'Mahony, Daniel J.
; TITLE OF INVENTION: MEMBRANE TRANSLOCATING PEPTIDE DRUG DELIVERY SYSTEM
; FILE REFERENCE: E1067/20018
; CURRENT APPLICATION NUMBER: US/09/671,089
; CURRENT FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/156,246
; PRIOR FILING DATE: 1999-09-27
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 2
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: membrane translocating peptide
; NAME/KEY: MOD_RES
; LOCATION: (15)..(15)
; OTHER INFORMATION: linked to FITC-LC
US-09-671-089-2
```

```
Query Match 67.1%; Score 51; DB 4; Length 15;
Best Local Similarity 92.3%; Pred. No. 0.071;
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
Qy 5 KRAAAVLLP-VLLA 17
Db 1 KKAHAVLLP-VLLA 13
```

```
RESULT 9
US-09-671-089-4
; Sequence 4, Application US/09671089
; Patent No. 6780846
; GENERAL INFORMATION:
; APPLICANT: O'Mahony, Daniel J.
; TITLE OF INVENTION: MEMBRANE TRANSLOCATING PEPTIDE DRUG DELIVERY SYSTEM
; FILE REFERENCE: E1067/20018
; CURRENT APPLICATION NUMBER: US/09/671,089
; CURRENT FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/156,246
; PRIOR FILING DATE: 1999-09-27
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: Patentin version 3.1
```

; SEQ ID NO 4  
; LENGTH: 19  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: membrane translocating peptide  
US-09-671-089-4

Query Match 67.1%; Score 51; DB 4; Length 19;  
Best Local Similarity 92.3%; Pred. No. 0.09;  
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 KAAAVLLPVLLA 17  
| | | | | | | | | |  
Db 1 KCAAVLLPVLLA 13

RESULT 10  
US-09-671-089-7  
; Sequence 7, Application US/09671089  
; Patent No. 6780846  
; GENERAL INFORMATION:  
; APPLICANT: O'Mahony, Daniel J.  
; APPLICANT: Lambkin, Imelda J.  
; TITLE OF INVENTION: MEMBRANE TRANSLOCATING PEPTIDE DRUG DELIVERY SYSTEM  
; FILE REFERENCE: E1067/20018  
; CURRENT APPLICATION NUMBER: US/09/671,089  
; CURRENT FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: 60/156,246  
; PRIOR FILING DATE: 1999-09-27  
; NUMBER OF SEQ ID NOS: 59  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 7  
; LENGTH: 14  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: membrane translocating peptide, cyclic internal  
US-09-671-089-7

Query Match 61.8%; Score 47; DB 4; Length 14;  
Best Local Similarity 84.6%; Pred. No. 0.28;  
Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 KAAAVLLPVLLA 17  
| | | | | | | | | |  
Db 1 KCAAVLLPVLLA 13

RESULT 11  
US-09-671-089-58  
; Sequence 58, Application US/09671089  
; Patent No. 6780846  
; GENERAL INFORMATION:  
; APPLICANT: O'Mahony, Daniel J.  
; APPLICANT: Lambkin, Imelda J.  
; TITLE OF INVENTION: MEMBRANE TRANSLOCATING PEPTIDE DRUG DELIVERY SYSTEM  
; FILE REFERENCE: E1067/20018  
; CURRENT APPLICATION NUMBER: US/09/671,089  
; CURRENT FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: 60/156,246  
; PRIOR FILING DATE: 1999-09-27  
; NUMBER OF SEQ ID NOS: 59  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 58  
; LENGTH: 14  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: membrane translocating peptide, cyclic internal  
US-09-671-089-58

Query Match 61.8%; Score 47; DB 4; Length 14;

Best Local Similarity 84.6%; Pred. No. 0.28;  
Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 5 KAAAVLLPVLLA 17  
| | | | | | | | | |  
Db 1 KCAAVLLPVLLA 13

RESULT 12  
US-09-671-089-5  
; Sequence 5, Application US/09671089  
; Patent No. 6780846  
; GENERAL INFORMATION:  
; APPLICANT: O'Mahony, Daniel J.  
; APPLICANT: Lambkin, Imelda J.  
; TITLE OF INVENTION: MEMBRANE TRANSLOCATING PEPTIDE DRUG DELIVERY SYSTEM  
; FILE REFERENCE: E1067/20018  
; CURRENT APPLICATION NUMBER: US/09/671,089  
; CURRENT FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: 60/156,246  
; PRIOR FILING DATE: 1999-09-27  
; NUMBER OF SEQ ID NOS: 59  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 5  
; LENGTH: 16  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: membrane translocating peptide, cyclic  
US-09-671-089-5

Query Match 61.8%; Score 47; DB 4; Length 16;  
Best Local Similarity 84.6%; Pred. No. 0.32;  
Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 KAAAVLLPVLLA 17  
| | | | | | | | | |  
Db 1 KCAAVLLPVLLA 13

RESULT 13  
US-08-928-958-2  
; Sequence 2, Application US/08928958  
; Patent No. 5877282  
; GENERAL INFORMATION:  
; APPLICANT: NADLER, STEVEN G.  
; APPLICANT: CLEAVELAND, JEFFREY S.  
; APPLICANT: BLAKE, JAMES  
; APPLICANT: HAFAR, OMAR K.  
; TITLE OF INVENTION: PEPTIDE INHIBITORS OF NUCLEAR PROTEIN  
; TITLE OF INVENTION: TRANSLOCATION HAVING NUCLEAR LOCALIZATION SEQUENCES AND  
; TITLE OF INVENTION: METHODS OF USE THEREOF  
; NUMBER OF SEQUENCES: 24  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: ROBINS & ASSOCIATES  
; STREET: 90 MIDDLEFIELD ROAD, SUITE 200  
; CITY: MENLO PARK  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94025  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/928,958  
; FILING DATE: 12-SEP-1997  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/026978  
; FILING DATE: 20-SEP-1996  
; ATTORNEY/AGENT INFORMATION:

; NAME: ROBINS, ROBERTA L.  
; REGISTRATION NUMBER: 33,208  
; REFERENCE/DOCKET NUMBER: 5998-0019  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (650) 325-7812  
; TELEFAX: (650) 325-7823  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 26 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-928-958-2

Query Match 61.8%; Score 47; DB 2; Length 26;  
Best Local Similarity 70.6%; Pred. No. 0.53;  
Matches 12; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 KKKRKAATAVLLPVLLA 17  
||| ||| | : |||  
DB 1 KKKYKAVALLPVLLA 17

RESULT 14  
US-09-072-429-2  
; Sequence 2, Application US/09072429  
; Patent No. 5962415  
; GENERAL INFORMATION:  
; APPLICANT: Nadler, Steven G.  
; TITLE OF INVENTION: COMPOSITIONS COMPRISING A PEPTIDE  
; TITLE OF INVENTION: INHIBITOR OF NUCLEAR PROTEIN TRANSLOCATION AND AN  
; TITLE OF INVENTION: IMMUNOSUPPRESSANT AND METHODS OF USE THEREOF  
; NUMBER OF SEQUENCES: 24  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Bristol-Myers Squibb Company  
; STREET: P.O. Box 4000  
; CITY: Princeton  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 08543-4000  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/072,429  
; FILING DATE: 04-MAY-1998  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Klein, Christopher A.  
; REGISTRATION NUMBER: 34,363  
; REFERENCE/DOCKET NUMBER: 090141b  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (609) 252-3714  
; TELEFAX: (609) 252-4526  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 26 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-09-072-429-2

Query Match 61.8%; Score 47; DB 2; Length 26;  
Best Local Similarity 70.6%; Pred. No. 0.53;  
Matches 12; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 KKKRKAATAVLLPVLLA 17  
||| ||| | : |||  
DB 1 KKKYKAVALLPVLLA 17

RESULT 15  
US-10-116-288A-22  
; Sequence 22, Application US/10116288A  
; Patent No. 6780843  
; GENERAL INFORMATION:  
; APPLICANT: Lin, Yao-Zhong  
; APPLICANT: Donahue, John P.  
; APPLICANT: Rojas, Mauricio  
; APPLICANT: Tan, Zhongjia  
; TITLE OF INVENTION: "Sequence and Method for Genetic Engineering of  
; Patent No. 6780843  
; TITLE OF INVENTION: Proteins with Cell Membrane Translocating Activity"  
; FILE REFERENCE: 22000.009703  
; CURRENT APPLICATION NUMBER: US/10/116,288A  
; CURRENT FILING DATE: 2002-04-04  
; PRIOR APPLICATION NUMBER: 09/562,868  
; PRIOR FILING DATE: 2000-05-01  
; PRIOR APPLICATION NUMBER: 09/186,170  
; PRIOR FILING DATE: 1998-11-04  
; PRIOR APPLICATION NUMBER: 60/080,083  
; PRIOR FILING DATE: 1998-03-31  
; NUMBER OF SEQ ID NOS: 23  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 22  
; LENGTH: 19  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Amino acid  
; OTHER INFORMATION: sequence of peptide which transports proteins  
; OTHER INFORMATION: through the cell membrane into the cell.  
US-10-116-288A-22

Query Match 57.9%; Score 44; DB 4; Length 19;  
Best Local Similarity 78.6%; Pred. No. 1.2;  
Matches 11; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 RKAAAVLLPVLLA 17  
| ||| ||| |||  
DB 4 RGIPAAVLLPVLLA 17

RESULT 16  
US-10-116-288A-20  
; Sequence 20, Application US/10116288A  
; Patent No. 6780843  
; GENERAL INFORMATION:  
; APPLICANT: Lin, Yao-Zhong  
; APPLICANT: Donahue, John P.  
; APPLICANT: Rojas, Mauricio  
; APPLICANT: Tan, Zhongjia  
; TITLE OF INVENTION: "Sequence and Method for Genetic Engineering of  
; Patent No. 6780843  
; TITLE OF INVENTION: Proteins with Cell Membrane Translocating Activity"  
; FILE REFERENCE: 22000.009703  
; CURRENT APPLICATION NUMBER: US/10/116,288A  
; CURRENT FILING DATE: 2002-04-04  
; PRIOR APPLICATION NUMBER: 09/562,868  
; PRIOR FILING DATE: 2000-05-01  
; PRIOR APPLICATION NUMBER: 09/186,170  
; PRIOR FILING DATE: 1998-11-04  
; PRIOR APPLICATION NUMBER: 60/080,083  
; PRIOR FILING DATE: 1998-03-31  
; NUMBER OF SEQ ID NOS: 23  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 20  
; LENGTH: 25  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Amino acid



; OTHER INFORMATION: sequence of peptide which transports proteins  
; OTHER INFORMATION: through the cell membrane into the cell.  
US-10-116-288A-20

Query Match 57.9%; Score 44; DB 4; Length 25;  
Best Local Similarity 76.3%; Pred. No. 1.5; Mismatches 2; Indels 0; Gaps 0;  
Matches 10; Conservative 1;

QY 5 KAAAVLLPVLLA 17  
: |||||  
Db 4 RGIAVLLPVLLA 16

RESULT 17  
US-09-186-170-4  
; Sequence 4, Application US/09186170  
; Patent No. 6248558  
; GENERAL INFORMATION:  
; APPLICANT: Lin, Yao-Zhong  
; APPLICANT: Donahue, John P.  
; APPLICANT: Rojas, Mauricio  
; APPLICANT: Tan, Zhongjia  
; TITLE OF INVENTION: "Sequence and Method for Genetic Engineering of  
; Patent No. 6248558  
; TITLE OF INVENTION: Proteins with Cell Membrane Translocating Activity"  
; FILE REFERENCE: VU9841  
; CURRENT APPLICATION NUMBER: US/09/186,170  
; CURRENT FILING DATE: 1998-11-04  
; EARLIER APPLICATION NUMBER: 60/080,083  
; EARLIER FILING DATE: 1998-03-31  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 4  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Amino acid  
; OTHER INFORMATION: sequence of peptide which transports proteins  
; OTHER INFORMATION: through the cell membrane into the cell.  
; NAME/KEY: PEPTIDE  
; LOCATION: (1)..(10)  
; JOURNAL: Nature Biotechnology  
; VOLUME: 16  
; ISSUE: April  
; PAGES: 370-375  
; DATE: 1998-04-01  
; RELEVANT RESIDUES: 1 TO 12  
US-09-186-170-4

Query Match 56.6%; Score 43; DB 3; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.87;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 AAVLLPVLLA 17  
: |||||  
Db 1 AAVLLPVLLA 10

RESULT 18  
US-09-562-868-4  
; Sequence 4, Application US/09562868  
; Patent No. 6432680  
; GENERAL INFORMATION:  
; APPLICANT: Lin, Yao-Zhong  
; APPLICANT: Donahue, John P.  
; APPLICANT: Rojas, Mauricio  
; APPLICANT: Tan, Zhongjia  
; TITLE OF INVENTION: "Sequence and Method for Genetic Engineering of

Query Match 56.6%; Score 43; DB 3; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.87;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 AAVLLPVLLA 17  
: |||||  
Db 1 AAVLLPVLLA 10

RESULT 19  
US-10-116-288A-4  
; Sequence 4, Application US/10116288A  
; Patent No. 6780843  
; GENERAL INFORMATION:  
; APPLICANT: Lin, Yao-Zhong  
; APPLICANT: Donahue, John P.  
; APPLICANT: Rojas, Mauricio  
; APPLICANT: Tan, Zhongjia  
; TITLE OF INVENTION: "Sequence and Method for Genetic Engineering of  
; Patent No. 6780843  
; TITLE OF INVENTION: Proteins with Cell Membrane Translocating Activity"  
; FILE REFERENCE: 22000.009703  
; CURRENT APPLICATION NUMBER: US/10/116,288A  
; CURRENT FILING DATE: 2002-04-04  
; PRIOR APPLICATION NUMBER: 09/562,868  
; PRIOR FILING DATE: 2000-05-01  
; PRIOR APPLICATION NUMBER: 09/186,170  
; PRIOR FILING DATE: 1998-11-04  
; PRIOR APPLICATION NUMBER: 60/080,083  
; PRIOR FILING DATE: 1998-03-31  
; NUMBER OF SEQ ID NOS: 23  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 4  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Amino acid  
; OTHER INFORMATION: sequence of peptide which transports proteins  
; OTHER INFORMATION: through the cell membrane into the cell.  
; NAME/KEY: PEPTIDE  
; LOCATION: (1)..(10)  
; JOURNAL: Nature Biotechnology  
; VOLUME: 16  
; ISSUE: April  
; PAGES: 370-375  
; DATE: 1998-04-01  
; RELEVANT RESIDUES: 1 TO 12  
US-09-186-170-4

; Patent No. 6432680  
; TITLE OF INVENTION: Proteins with Cell Membrane Translocating Activity"  
; FILE REFERENCE: 22000.009702  
; CURRENT APPLICATION NUMBER: US/09/562,868  
; CURRENT FILING DATE: 2000-05-01  
; PRIOR APPLICATION NUMBER: 60/080,083  
; PRIOR FILING DATE: 1998-03-31  
; PRIOR APPLICATION NUMBER: 09/186,170  
; PRIOR FILING DATE: 1998-11-04  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 4  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Amino acid  
; OTHER INFORMATION: sequence of peptide which transports proteins  
; OTHER INFORMATION: through the cell membrane into the cell.  
; NAME/KEY: PEPTIDE  
; LOCATION: (1)..(10)  
; PUBLICATION INFORMATION:  
; AUTHORS: Rojas, M. et al.  
; TITLE: "Genetic Engineering of Proteins with Cell Membrane  
; TITLE: "Permeability"  
; JOURNAL: Nature Biotechnology  
; VOLUME: 16  
; ISSUE: April  
; PAGES: 370-375  
; DATE: 1998-04-01  
; SOFTWARE: PatentIn Ver. 2.0  
US-09-562-868-4

Query Match 56.6%; Score 43; DB 4; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.87;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 AAVLLPVLLA 17  
: |||||  
Db 1 AAVLLPVLLA 10

RESULT 19  
US-10-116-288A-4  
; Sequence 4, Application US/10116288A  
; Patent No. 6780843  
; GENERAL INFORMATION:  
; APPLICANT: Lin, Yao-Zhong  
; APPLICANT: Donahue, John P.  
; APPLICANT: Rojas, Mauricio  
; APPLICANT: Tan, Zhongjia  
; TITLE OF INVENTION: "Sequence and Method for Genetic Engineering of  
; Patent No. 6780843  
; TITLE OF INVENTION: Proteins with Cell Membrane Translocating Activity"  
; FILE REFERENCE: 22000.009703  
; CURRENT APPLICATION NUMBER: US/10/116,288A  
; CURRENT FILING DATE: 2002-04-04  
; PRIOR APPLICATION NUMBER: 09/562,868  
; PRIOR FILING DATE: 2000-05-01  
; PRIOR APPLICATION NUMBER: 09/186,170  
; PRIOR FILING DATE: 1998-11-04  
; PRIOR APPLICATION NUMBER: 60/080,083  
; PRIOR FILING DATE: 1998-03-31  
; NUMBER OF SEQ ID NOS: 23  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 4  
; LENGTH: 10  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Amino acid  
; OTHER INFORMATION: sequence of peptide which transports proteins  
; OTHER INFORMATION: through the cell membrane into the cell.  
; NAME/KEY: PEPTIDE  
; LOCATION: (1)..(10)  
; JOURNAL: Nature Biotechnology  
; VOLUME: 16  
; ISSUE: April  
; PAGES: 370-375  
; DATE: 1998-04-01  
; RELEVANT RESIDUES: 1 TO 12  
US-09-186-170-4

; NAME/KEY: PEPTIDE  
; LOCATION: (1)..(10)  
; PUBLICATION INFORMATION:  
; AUTHORS: Rojas, M. et al.  
; TITLE: "Genetic Engineering of Proteins with Cell Membrane  
; TITLE: Permeability"  
; JOURNAL: Nature Biotechnology  
; VOLUME: 16  
; ISSUE: April  
; PAGES: 370-375  
; DATE: 1998-04-01  
US-10-116-288A-4

Query Match 56.6%; Score 43; DB 4; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.87;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 8 AAVLLPVLLA 17  
| | | | | | | | | |  
Db 1 AAVLLPVLLA 10

RESULT 20  
US-09-186-170-5  
; Sequence 5, Application US/09186170  
; Patent No. 6248558  
; GENERAL INFORMATION:  
; APPLICANT: Lin, Yao-Zhong  
; APPLICANT: Donahue, John P.  
; APPLICANT: Rojas, Mauricio  
; APPLICANT: Tan, Zhongjia  
; TITLE OF INVENTION: "Sequence and Method for Genetic Engineering of  
; Patent No. 6248558  
; TITLE OF INVENTION: "Sequence and Method for Genetic Engineering of  
; FILE REFERENCE: VU9841  
; CURRENT APPLICATION NUMBER: US/09/186,170  
; CURRENT FILING DATE: 1998-11-04  
; EARLIER APPLICATION NUMBER: 60/080,083  
; EARLIER FILING DATE: 1998-03-31  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 5  
; LENGTH: 11  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Amino acid  
; OTHER INFORMATION: sequence of peptide which transports proteins  
; OTHER INFORMATION: through the cell membrane into the cell.  
; FEATURE:  
; NAME/KEY: PEPTIDE  
; LOCATION: (1)..(11)  
; PUBLICATION INFORMATION:  
; TITLE: "Genetic Engineering of Proteins with Cell Membrane  
; TITLE: Permeability"  
; JOURNAL: Nature Biotechnology  
; VOLUME: 16  
; ISSUE: April  
; PAGES: 370-375  
; DATE: 1998-04-01  
; RELEVANT RESIDUES: 1 TO 12  
US-09-186-170-5

Query Match 56.6%; Score 43; DB 3; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.96;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 8 AAVLLPVLLA 17  
| | | | | | | | | |  
Db 1 AAVLLPVLLA 10

RESULT 21

US-09-562-868-5  
; Sequence 5, Application US/09562868  
; Patent No. 6432680  
; GENERAL INFORMATION:  
; APPLICANT: Lin, Yao-Zhong  
; APPLICANT: Donahue, John P.  
; APPLICANT: Rojas, Mauricio  
; APPLICANT: Tan, Zhongjia  
; TITLE OF INVENTION: "Sequence and Method for Genetic Engineering of  
; Patent No. 6432680  
; TITLE OF INVENTION: Proteins with Cell Membrane Translocating Activity"  
; FILE REFERENCE: 22000.0097U2  
; CURRENT APPLICATION NUMBER: US/09/562,868  
; CURRENT FILING DATE: 2000-05-01  
; PRIOR APPLICATION NUMBER: 60/080,083  
; PRIOR FILING DATE: 1998-03-31  
; PRIOR APPLICATION NUMBER: 09/186,170  
; PRIOR FILING DATE: 1998-11-04  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 5  
; LENGTH: 11  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Amino acid  
; OTHER INFORMATION: sequence of peptide which transports proteins  
; OTHER INFORMATION: through the cell membrane into the cell.  
; NAME/KEY: PEPTIDE  
; LOCATION: (1)..(11)  
; PUBLICATION INFORMATION:  
; AUTHORS: Rojas, M. et al.  
; TITLE: "Genetic Engineering of Proteins with Cell Membrane  
; TITLE: Permeability"  
; JOURNAL: Nature Biotechnology  
; VOLUME: 16  
; ISSUE: April  
; PAGES: 370-375  
; DATE: 1998-04-01  
US-09-562-868-5

Query Match 56.6%; Score 43; DB 4; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.96;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 8 AAVLLPVLLA 17  
| | | | | | | | | |  
Db 1 AAVLLPVLLA 10

RESULT 22  
US-10-116-288A-5  
; Sequence 5, Application US/10116288A  
; Patent No. 6780843  
; GENERAL INFORMATION:  
; APPLICANT: Lin, Yao-Zhong  
; APPLICANT: Donahue, John P.  
; APPLICANT: Rojas, Mauricio  
; APPLICANT: Tan, Zhongjia  
; TITLE OF INVENTION: "Sequence and Method for Genetic Engineering of  
; Patent No. 6780843  
; TITLE OF INVENTION: Proteins with Cell Membrane Translocating Activity"  
; FILE REFERENCE: 22000.0097U3  
; CURRENT APPLICATION NUMBER: US/10/116,288A  
; CURRENT FILING DATE: 2002-04-04  
; PRIOR APPLICATION NUMBER: 09/562,868  
; PRIOR FILING DATE: 2000-05-01  
; PRIOR APPLICATION NUMBER: 09/186,170  
; PRIOR FILING DATE: 1998-11-04  
; PRIOR APPLICATION NUMBER: 60/080,083  
; PRIOR FILING DATE: 1998-03-31  
; NUMBER OF SEQ ID NOS: 23  
; SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 5  
LENGTH: 11  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Amino acid  
OTHER INFORMATION: sequence of peptide which transports proteins  
OTHER INFORMATION: through the cell membrane into the cell.  
FEATURE:  
NAME/KEY: PEPTIDE  
LOCATION: (1)..(11)  
PUBLICATION INFORMATION:  
AUTHORS: Rojas, M. et al.  
TITLE: "Genetic Engineering of Proteins with Cell Membrane  
JOURNAL: Nature Biotechnology  
VOLUME: 16  
ISSUE: April  
PAGES: 370-375  
DATE: 1998-04-01  
US-10-116-288A-5.

Query Match 56.6%; Score 43; DB 4; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.96;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 AAVLLPVLLA 17  
|||||

Db 1 AAVLLPVLLA 10

RESULT 23  
US-09-671-089-15  
Sequence 15, Application US/09671089  
Patent No. 6780846  
GENERAL INFORMATION:  
APPLICANT: O'Mahony, Daniel J.  
APPLICANT: Lambkin, Imelda J.  
TITLE OF INVENTION: MEMBRANE TRANSLOCATING PEPTIDE DRUG DELIVERY SYSTEM  
FILE REFERENCE: E1067/20018  
CURRENT APPLICATION NUMBER: US/09/671,089  
CURRENT FILING DATE: 2000-09-27  
PRIOR APPLICATION NUMBER: 60/156,246  
PRIOR FILING DATE: 1999-09-27  
NUMBER OF SEQ ID NOS: 59  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 15  
LENGTH: 11  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: membrane translocating peptide  
US-09-671-089-15

Query Match 56.6%; Score 43; DB 4; Length 11;  
Best Local Similarity 100.0%; Pred. No. 0.96;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 AAVLLPVLLA 17  
|||||

Db 1 AAVLLPVLLA 10

RESULT 24  
US-09-186-170-1  
Sequence 1, Application US/09186170  
Patent No. 6248558  
GENERAL INFORMATION:  
APPLICANT: Lin, Yao-Zhong  
APPLICANT: Donahue, John P.  
APPLICANT: Rojas, Mauricio  
APPLICANT: Tan, Zhongjia  
TITLE OF INVENTION: "Sequence and Method for Genetic Engineering of

Patent No. 6248558  
TITLE OF INVENTION: Proteins with Cell Membrane Translocating Activity"  
FILE REFERENCE: VU9841  
CURRENT APPLICATION NUMBER: US/09/186,170  
CURRENT FILING DATE: 1998-11-04  
EARLIER APPLICATION NUMBER: 60/080,083  
EARLIER FILING DATE: 1998-03-31  
NUMBER OF SEQ ID NOS: 18  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 1  
LENGTH: 12  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Amino acid  
OTHER INFORMATION: sequence of peptide which transports proteins  
OTHER INFORMATION: through the cell membrane into the cell.  
FEATURE:  
NAME/KEY: PEPTIDE  
LOCATION: (1)..(12)  
PUBLICATION INFORMATION:  
TITLE: "Genetic Engineering of Proteins with Cell Membrane  
JOURNAL: Nature Biotechnology  
VOLUME: 16  
ISSUE: April  
PAGES: 370-375  
DATE: 1998-04-01  
RELEVANT RESIDUES: 1 TO 12  
US-09-186-170-1

Query Match 56.6%; Score 43; DB 3; Length 12;  
Best Local Similarity 100.0%; Pred. No. 1;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 AAVLLPVLLA 17  
|||||

Db 1 AAVLLPVLLA 10

RESULT 25  
US-09-562-868-1  
Sequence 1, Application US/09562868  
Patent No. 6432680  
GENERAL INFORMATION:  
APPLICANT: Lin, Yao-Zhong  
APPLICANT: Donahue, John P.  
APPLICANT: Rojas, Mauricio  
APPLICANT: Tan, Zhongjia  
TITLE OF INVENTION: "Sequence and Method for Genetic Engineering of  
Patent No. 6432680  
TITLE OF INVENTION: Proteins with Cell Membrane Translocating Activity"  
FILE REFERENCE: 22000.009702  
CURRENT APPLICATION NUMBER: US/09/562,868  
CURRENT FILING DATE: 2000-05-01  
PRIOR APPLICATION NUMBER: 60/080,083  
PRIOR FILING DATE: 1998-03-31  
PRIOR APPLICATION NUMBER: 09/186,170  
PRIOR FILING DATE: 1998-11-04  
NUMBER OF SEQ ID NOS: 18  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 1  
LENGTH: 12  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Amino acid  
OTHER INFORMATION: sequence of peptide which transports proteins  
OTHER INFORMATION: through the cell membrane into the cell.  
NAME/KEY: PEPTIDE  
LOCATION: (1)..(12)  
PUBLICATION INFORMATION:  
AUTHORS: Rojas, M. et al.

; TITLE: "Genetic Engineering of Proteins with Cell Membrane  
; TITLE: Permeability"  
; JOURNAL: Nature Biotechnology  
; VOLUME: 16  
; ISSUE: April  
; PAGES: 370-375  
; DATE: 1998-04-01  
US-09-562-868-1

Query Match 56.6%; Score 43; DB 4; Length 12;  
Best Local Similarity 100.0%; Pred. No. 1;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 AAVLLPVLLA 17  
| | | | | | | | | |  
Db 1 AAVLLPVLLA 10

RESULT 26  
US-09-997-465B-4  
; Sequence 4, Application US/09997465B  
; Patent No. 6673574  
; GENERAL INFORMATION:  
; APPLICANT: Stern, William  
; APPLICANT: Mehta, No. 6673574er M.  
; TITLE OF INVENTION: IMPROVED ORAL DELIVERY OF PEPTIDES USING ENZYME-CLEAVABLE MEMBRANES  
; FILE REFERENCE: P/546-247  
; CURRENT APPLICATION NUMBER: US/09/997,465B  
; CURRENT FILING DATE: 2001-11-29  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 4  
; LENGTH: 12  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-997-465B-4

Query Match 56.6%; Score 43; DB 4; Length 12;  
Best Local Similarity 100.0%; Pred. No. 1;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 AAVLLPVLLA 17  
| | | | | | | | | |  
Db 1 AAVLLPVLLA 10

RESULT 27  
US-10-083-889-8  
; Sequence 8, Application US/10083889  
; Patent No. 6673894  
; GENERAL INFORMATION:  
; APPLICANT: Zahner, Joseph E.  
; TITLE OF INVENTION: Inhibitor of cell proliferation and methods of use thereof.  
; FILE REFERENCE: 16850-7331  
; CURRENT APPLICATION NUMBER: US/10/083,889  
; CURRENT FILING DATE: 2002-02-27  
; PRIOR APPLICATION NUMBER: US 60/271,798  
; PRIOR FILING DATE: 2001-02-27  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: Microsoft Word 97  
; SEQ ID NO 8  
; LENGTH: 12  
; TYPE: PRT  
; ORGANISM: Mammalian  
US-10-083-889-8

Query Match 56.6%; Score 43; DB 4; Length 12;  
Best Local Similarity 100.0%; Pred. No. 1;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 AAVLLPVLLA 17

Db 1 AAVLLPVLLA 10

RESULT 28  
US-10-116-288A-1  
; Sequence 1, Application US/10116288A  
; Patent No. 6780843  
; GENERAL INFORMATION:  
; APPLICANT: Lin, Yao-Zhong  
; APPLICANT: Donahue, John P.  
; APPLICANT: Rojas, Mauricio  
; APPLICANT: Tan, Zhongjia  
; TITLE OF INVENTION: "Sequence and Method for Genetic Engineering of Proteins with Cell Membrane Translocating Activity"  
; Patent No. 6780843  
; TITLE OF INVENTION: Proteins with Cell Membrane Translocating Activity"  
; FILE REFERENCE: 22000.0097U3  
; CURRENT APPLICATION NUMBER: US/10/116,288A  
; CURRENT FILING DATE: 2002-04-04  
; PRIOR APPLICATION NUMBER: 09/562,868  
; PRIOR FILING DATE: 2000-05-01  
; PRIOR APPLICATION NUMBER: 09/186,170  
; PRIOR FILING DATE: 1998-11-04  
; PRIOR APPLICATION NUMBER: 60/080,083  
; PRIOR FILING DATE: 1998-03-31  
; NUMBER OF SEQ ID NOS: 23  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1  
; LENGTH: 12  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: Amino acid sequence of peptide which transports proteins through the cell membrane into the cell.  
; OTHER INFORMATION: through the cell membrane into the cell.  
; FEATURE:  
; NAME/KEY: PEPTIDE  
; LOCATION: (1)..(12)  
; PUBLICATION INFORMATION:  
; AUTHORS: Rojas, M. et al.  
; TITLE: "Genetic Engineering of Proteins with Cell Membrane Permeability"  
; JOURNAL: Nature Biotechnology  
; VOLUME: 16  
; ISSUE: April  
; PAGES: 370-375  
; DATE: 1998-04-01  
US-10-116-288A-1

Query Match 56.6%; Score 43; DB 4; Length 12;  
Best Local Similarity 100.0%; Pred. No. 1;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 AAVLLPVLLA 17  
| | | | | | | | | |  
Db 1 AAVLLPVLLA 10

RESULT 29  
US-09-671-089-1  
; Sequence 1, Application US/09671089  
; Patent No. 6780846  
; GENERAL INFORMATION:  
; APPLICANT: O'Mahony, Daniel J.  
; APPLICANT: Lambkin, Imelda J.  
; TITLE OF INVENTION: MEMBRANE TRANSLOCATING PEPTIDE DRUG DELIVERY SYSTEM  
; FILE REFERENCE: E1067/20018  
; CURRENT APPLICATION NUMBER: US/09/671,089  
; CURRENT FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: 60/156,246  
; PRIOR FILING DATE: 1999-09-27  
; NUMBER OF SEQ ID NOS: 59  
; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 1  
; LENGTH: 12  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: membrane translocating peptide  
US-09-671-089-1

Query Match 56.6%; Score 43; DB 4; Length 12;  
Best Local Similarity 100.0%; Pred. No. 1;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 AAVLLPVLLA 17  
| | | | | | | | | |  
DB 1 AAVLLPVLLA 10

RESULT 30  
US-09-671-089-14  
; Sequence 14, Application US/09671089  
; Patent No. 6780846  
; GENERAL INFORMATION:  
; APPLICANT: O'Mahony, Daniel J.  
; IMELDA J.  
; TITLE OF INVENTION: MEMBRANE TRANSLOCATING PEPTIDE DRUG DELIVERY SYSTEM  
; FILE REFERENCE: E1067/20018  
; CURRENT APPLICATION NUMBER: US/09/671,089  
; CURRENT FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: 60/156,246  
; PRIOR FILING DATE: 1999-09-27  
; NUMBER OF SEQ ID NOS: 59  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 14  
; LENGTH: 12  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: membrane translocating peptide  
US-09-671-089-14

Query Match 56.6%; Score 43; DB 4; Length 12;  
Best Local Similarity 100.0%; Pred. No. 1;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 AAVLLPVLLA 17  
| | | | | | | | | |  
DB 1 AAVLLPVLLA 10

RESULT 31  
US-09-671-089-16  
; Sequence 16, Application US/09671089  
; Patent No. 6780846  
; GENERAL INFORMATION:  
; APPLICANT: O'Mahony, Daniel J.  
; IMELDA J.  
; TITLE OF INVENTION: MEMBRANE TRANSLOCATING PEPTIDE DRUG DELIVERY SYSTEM  
; FILE REFERENCE: E1067/20018  
; CURRENT APPLICATION NUMBER: US/09/671,089  
; CURRENT FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: 60/156,246  
; PRIOR FILING DATE: 1999-09-27  
; NUMBER OF SEQ ID NOS: 59  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 16  
; LENGTH: 12  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: membrane translocating peptide  
US-09-671-089-16

Query Match 56.6%; Score 43; DB 4; Length 12;

Best Local Similarity 100.0%; Pred. No. 1;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 8 AAVLLPVLLA 17  
| | | | | | | | | |  
DB 3 AAVLLPVLLA 12

RESULT 32  
US-09-671-089-59  
; Sequence 59, Application US/09671089  
; Patent No. 6780846  
; GENERAL INFORMATION:  
; APPLICANT: O'Mahony, Daniel J.  
; IMELDA J.  
; TITLE OF INVENTION: MEMBRANE TRANSLOCATING PEPTIDE DRUG DELIVERY SYSTEM  
; FILE REFERENCE: E1067/20018  
; CURRENT APPLICATION NUMBER: US/09/671,089  
; CURRENT FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: 60/156,246  
; PRIOR FILING DATE: 1999-09-27  
; NUMBER OF SEQ ID NOS: 59  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 59  
; LENGTH: 12  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: membrane translocating peptide  
US-09-671-089-59

Query Match 56.6%; Score 43; DB 4; Length 12;  
Best Local Similarity 100.0%; Pred. No. 1;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 AAVLLPVLLA 17  
| | | | | | | | | |  
DB 3 AAVLLPVLLA 12

RESULT 33  
US-10-144-549-19  
; Sequence 19, Application US/10144549  
; Patent No. 6835810  
; GENERAL INFORMATION:  
; APPLICANT: Geneshuttle Biopharm, Inc.  
; Hwu, Paul L.  
; TITLE OF INVENTION: A NEW FUSION PROTEIN FOR USE AS VECTOR  
; FILE REFERENCE: MBHB 02-340  
; CURRENT APPLICATION NUMBER: US/10/144,549  
; CURRENT FILING DATE: 2002-05-13  
; NUMBER OF SEQ ID NOS: 31  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 19  
; LENGTH: 12  
; TYPE: PRT  
; ORGANISM: Kaposi's sarcoma-associated herpesvirus  
; FEATURE:  
; NAME/KEY: MISC FEATURE  
; OTHER INFORMATION: Kaposi PGF signal sequence.  
US-10-144-549-19

Query Match 56.6%; Score 43; DB 4; Length 12;  
Best Local Similarity 100.0%; Pred. No. 1;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 8 AAVLLPVLLA 17  
| | | | | | | | | |  
DB 1 AAVLLPVLLA 10

RESULT 34  
US-09-671-089-6

```
; Sequence 6, Application US/09671089
; Patent No. 6780846
; GENERAL INFORMATION:
; APPLICANT: O'Mahony, Daniel J.
; APPLICANT: Lambkin, Imelda J.
; TITLE OF INVENTION: MEMBRANE TRANSLOCATING PEPTIDE DRUG DELIVERY SYSTEM
; FILE REFERENCE: E1067/20018
; CURRENT APPLICATION NUMBER: US/09/671,089
; CURRENT FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/156,246
; PRIOR FILING DATE: 1999-09-27
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: membrane translocating peptide, cyclic
US-09-671-089-6

Query Match      56.6%; Score 43; DB 4; Length 13;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      8 AAVLLPVLLA 17
DB      2 AAVLLPVLLA 11

RESULT 35
US-09-671-089-23
; Sequence 23, Application US/09671089
; Patent No. 6780846
; GENERAL INFORMATION:
; APPLICANT: O'Mahony, Daniel J.
; APPLICANT: Lambkin, Imelda J.
; TITLE OF INVENTION: MEMBRANE TRANSLOCATING PEPTIDE DRUG DELIVERY SYSTEM
; FILE REFERENCE: E1067/20018
; CURRENT APPLICATION NUMBER: US/09/671,089
; CURRENT FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/156,246
; PRIOR FILING DATE: 1999-09-27
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 23
; LENGTH: 17
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: membrane translocating peptide
US-09-671-089-23

Query Match      56.6%; Score 43; DB 4; Length 17;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      8 AAVLLPVLLA 17
DB      1 AAVLLPVLLA 10

RESULT 36
US-09-252-991A-20502
; Sequence 20502, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
```

```
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 20502
; LENGTH: 456
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-20502

Query Match      56.6%; Score 43; DB 4; Length 456;
Best Local Similarity 52.9%; Pred. No. 41;
Matches 9; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY      1 KKKRKAATAAALLPVLLA 17
DB      16 RRRERAGAVACPVLIA 32

RESULT 37
US-08-448-489-7
; Sequence 7, Application US/08448489
; Patent No. 6184022
; GENERAL INFORMATION:
; APPLICANT: SEIKI, Motoharu
; APPLICANT: SATO, Hiroshi
; APPLICANT: SHINAGAWA, Akira
; TITLE OF INVENTION: NOVEL METALLOPROTEINASE AND ENCODING DNA THEREFOR
; FILE REFERENCE: 55-290P
; CURRENT APPLICATION NUMBER: US/08/448,489
; CURRENT FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 30
; TYPE: PRT
; ORGANISM: Homo sapiens
US-08-448-489-7

Query Match      53.9%; Score 41; DB 3; Length 30;
Best Local Similarity 81.8%; Pred. No. 5.4;
Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY      6 AAAAALLPVLL 16
DB      6 SAAAVVLPVLL 16

RESULT 38
US-09-689-730-7
; Sequence 7, Application US/09689730
; Patent No. 6825024
; GENERAL INFORMATION:
; APPLICANT: SEIKI, Motoharu
; APPLICANT: SATO, Hiroshi
; APPLICANT: SHINAGAWA, Akira
; TITLE OF INVENTION: NOVEL METALLOPROTEINASE AND ENCODING DNA THEREFOR
; FILE REFERENCE: 55-290P
; CURRENT APPLICATION NUMBER: US/09/689,730
; CURRENT FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: US/08/448,489
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 30
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-689-730-7

Query Match      53.9%; Score 41; DB 4; Length 30;
Best Local Similarity 81.8%; Pred. No. 5.4;
```

Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 6 AAAAVALLPVLL 16  
:||||:|||||  
Db 6 SAAAVLPVLL 16

RESULT 39  
US-09-252-991A-19092  
; Sequence 19092, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,789  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 19092  
; LENGTH: 421  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-19092

Query Match 53.9%; Score 41; DB 4; Length 421;  
Best Local Similarity 50.0%; Pred. No. 78;  
Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 2 KKRKAAVALLPVLLA 17  
:::|||||:|:  
Db 48 RRRRAAGALRPVLLA 63

RESULT 40  
US-08-704-711A-1  
; Sequence 1, Application US/08704711A  
; Patent No. 6114159  
; GENERAL INFORMATION:  
; APPLICANT: WILL, Horst  
; APPLICANT: HINZMANN, Bernd  
; TITLE OF INVENTION: DNA SEQUENCES FOR MATRIX  
; TITLE OF INVENTION: METALLOPROTEASES, THEIR PRODUCTION AND USE  
; NUMBER OF SEQUENCES: 22  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 3000 K Street, N.W., Suite 500  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20007-5109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/704,711A  
; FILING DATE: 20-NOV-1996  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: WO PCT/DE95/00357  
; FILING DATE: 17-MAR-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: DE 4438838.1  
; FILING DATE: 21-OCT-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: DE 4409663.1  
; FILING DATE: 17-MAR-1994

ATTORNEY/AGENT INFORMATION:  
; NAME: GRANADOS, Patricia D.  
; REGISTRATION NUMBER: 33,683  
; REFERENCE/DOCKET NUMBER: 26083/124  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202)672-5300  
; TELEFAX: (202)672-5399  
; TELEX: 904136  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 579 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-08-704-711A-1

Query Match 53.9%; Score 41; DB 3; Length 579;  
Best Local Similarity 81.8%; Pred. No. 1.1e+02;  
Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 6 AAAAVALLPVLL 16  
:||||:|||||  
Db 535 SAAAVLPVLL 545

RESULT 41  
US-09-521-220-1  
; Sequence 1, Application US/09521220  
; Patent No. 6399348  
; GENERAL INFORMATION:  
; APPLICANT: WILL, Horst  
; APPLICANT: HINZMANN, Bernd  
; TITLE OF INVENTION: DNA SEQUENCES FOR MATRIX  
; NUMBER OF SEQUENCES: 22  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 3000 K Street, N.W., Suite 500  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20007-5109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/521,220  
; FILING DATE: 08-Mar-2000  
; CLASSIFICATION: <unknown>  
; 21-OCT-1994  
; 17-MAR-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/704,711  
; FILING DATE: <Unknown>  
; APPLICATION NUMBER: DE 4438838.1  
; FILING DATE: 21-OCT-1994  
; APPLICATION NUMBER: DE 4409663.1  
; FILING DATE: 17-MAR-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: GRANADOS, Patricia D.  
; REGISTRATION NUMBER: 33,683  
; REFERENCE/DOCKET NUMBER: 26083/124  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202)672-5300  
; TELEFAX: (202)672-5399  
; TELEX: 904136  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 579 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single

```

;
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-521-220-1
;
; Query Match 53.9%; Score 41; DB 3; Length 579;
; Best Local Similarity 81.8%; Pred. No. 1.1e+02;
; Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
;
QY 6 AAAAVLLPVLL 16
; :|||:||||
DB 535 SAAAVLPLVLL 545
;
;
RESULT 42
US-08-704-711A-2
; Sequence 2, Application US/08704711A
; Patent No. 6114159
; GENERAL INFORMATION:
; APPLICANT: WILL, Horet
; TITLE OF INVENTION: DNA SEQUENCES FOR MATRIX
; TITLE OF INVENTION: METALLOPROTEASES, THEIR PRODUCTION AND USE
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/704,711A
; FILING DATE: 20-NOV-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/DE95/00357
; FILING DATE: 17-MAR-1995
; PRIOR APPLICATION DATA: DE 4438838.1
; APPLICATION NUMBER: DE 4409663.1
; FILING DATE: 17-MAR-1994
; APPLICATION DATA:
; FILING DATE: 17-MAR-1994
; NAME: GRANADOS, Patricia D.
; REGISTRATION NUMBER: 33,683
; REFERENCE/DOCKET NUMBER: 26083/124
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 582 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
US-08-704-711A-2
;
; Query Match 53.9%; Score 41; DB 3; Length 582;
; Best Local Similarity 81.8%; Pred. No. 1.1e+02;
; Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
;
QY 6 AAAAVLLPVLL 16
; :|||:||||
DB 538 SAAAVLPLVLL 548
;
;
RESULT 43
US-08-448-489-1
; Sequence 1, Application US/08448489
; Patent No. 6184022
; GENERAL INFORMATION:
; APPLICANT: SEIKI, Motoharu
; APPLICANT: SATO, Hiroshi
; APPLICANT: SHINAGAWA, Akira
; TITLE OF INVENTION: NOVEL METALLOPROTEINASE AND ENCODING DNA THEREFOR
; FILE REFERENCE: 55-290P
; CURRENT APPLICATION NUMBER: US/08/448,489
; CURRENT FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 1
; LENGTH: 582
; TYPE: PRT
; ORGANISM: Homo sapiens
;
US-08-448-489-1
;
; Query Match 53.9%; Score 41; DB 3; Length 582;
; Best Local Similarity 81.8%; Pred. No. 1.1e+02;
; Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
;
QY 6 AAAAVLLPVLL 16
; :|||:||||
DB 538 SAAAVLPLVLL 548
;
;
RESULT 44
US-09-211-704A-9
; Sequence 9, Application US/09211704A
; Patent No. 6271014
; GENERAL INFORMATION:
; APPLICANT: de Saint-Vis, Blandine Marie
; APPLICANT: Fossiez, Francois
; APPLICANT: Caux, Christophe
; APPLICANT: Lebecque, Serge J.E.
; TITLE OF INVENTION: Mammalian Proteinases; Related Reagents
; TITLE OF INVENTION: and Methods
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DNAX Research Institute
; STREET: 901 California Avenue
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304-1104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/211,704A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/005,263
; FILING DATE: 09-JAN-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Ching, Edwin P.
; REGISTRATION NUMBER: 34,090
; REFERENCE/DOCKET NUMBER: SF0781K
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650)852-9196
; TELEFAX: (650)496-1200
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 582 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: linear
;

```



MOLECULE TYPE: peptide  
US-09-211-704A-9

Query Match 53.9%; Score 41; DB 3; Length 582;  
Best Local Similarity 81.8%; Pred. No. 1.1e+02;  
Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 6 AAAAVLLPVLL 16  
:||||:|||||  
Db 538 SAAAVLPLVLL 548

## RESULT 45

US-09-521-220-2  
; Sequence 2, Application US/09521220  
; Patent No. 6399348  
; GENERAL INFORMATION:  
; APPLICANT: WILL, Horst  
; TITLE OF INVENTION: DNA SEQUENCES FOR MATRIX  
; METALLOPROTEASES, THEIR PRODUCTION AND USE  
; NUMBER OF SEQUENCES: 22  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 3000 K Street, N.W., Suite 500  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20007-5109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA: US/09/521,220  
; APPLICATION NUMBER: US/09/521,220  
; FILING DATE: 08-Mar-2000  
; CLASSIFICATION: <Unknown>  
; 21-OCT-1994  
; 17-MAR-1994  
; PRIOR APPLICATION DATA: 08/704,711  
; FILING DATE: <Unknown>  
; APPLICATION NUMBER: DE 4438838.1  
; FILING DATE: 21-OCT-1994  
; APPLICATION NUMBER: DE 4409663.1  
; FILING DATE: 17-MAR-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: GRANADOS, Patricia D.  
; REGISTRATION NUMBER: 33,683  
; REFERENCE/DOCKET NUMBER: 26083/124  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202)672-5300  
; TELEFAX: (202)672-5399  
; TELEX: 904136  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 582 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:

Query Match 53.9%; Score 41; DB 3; Length 582;  
Best Local Similarity 81.8%; Pred. No. 1.1e+02;  
Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 6 AAAAVLLPVLL 16  
:||||:|||||  
Db 538 SAAAVLPLVLL 548

## RESULT 46

US-09-391-104-28  
; Sequence 28, Application US/093911104  
; Patent No. 6399371  
; GENERAL INFORMATION:  
; APPLICANT: Abbott Laboratories  
; APPLICANT: Falduto, Michael T.  
; APPLICANT: Magnuson, Scott R.  
; APPLICANT: Morgan, Douglas W.  
; TITLE OF INVENTION: HUMAN MATRIX METALLOPROTEASE GENE,  
; TITLE OF INVENTION: PROTEINS ENCODED THEREFROM AND METHODS  
; FILE REFERENCE: 6073.US.P1  
; CURRENT APPLICATION NUMBER: US/09/391,104  
; CURRENT FILING DATE: 1999-09-07  
; PRIOR APPLICATION NUMBER: US 08/814,394  
; PRIOR FILING DATE: 1997-03-11  
; NUMBER OF SEQ ID NOS: 35  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 28  
; LENGTH: 582  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-391-104-28

Query Match 53.9%; Score 41; DB 3; Length 582;  
Best Local Similarity 81.8%; Pred. No. 1.1e+02;  
Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 6 AAAAVLLPVLL 16  
:||||:|||||  
Db 538 SAAAVLPLVLL 548

## RESULT 47

US-09-919-497-84  
; Sequence 84, Application US/09919497  
; Patent No. 6773883  
; GENERAL INFORMATION:  
; APPLICANT: Mutter, George L.  
; TITLE OF INVENTION: PROGNOSTIC CLASSIFICATION OF ENDOMETRIAL CANCER  
; FILE REFERENCE: B0801/7225  
; CURRENT APPLICATION NUMBER: US/09/919,497  
; CURRENT FILING DATE: 2001-07-31  
; PRIOR APPLICATION NUMBER: US 60/221,735  
; PRIOR FILING DATE: 2000-07-31  
; NUMBER OF SEQ ID NOS: 100  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 84  
; LENGTH: 582  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-919-497-84

Query Match 53.9%; Score 41; DB 4; Length 582;  
Best Local Similarity 81.8%; Pred. No. 1.1e+02;  
Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 6 AAAAVLLPVLL 16  
:||||:|||||  
Db 538 SAAAVLPLVLL 548

## RESULT 48

US-09-689-730-1  
; Sequence 1, Application US/09689730  
; Patent No. 6825024  
; GENERAL INFORMATION:  
; APPLICANT: SEIKI, Motoharu  
; APPLICANT: SATO, Hiroshi  
; APPLICANT: SHINAGAWA, Akira  
; TITLE OF INVENTION: NOVEL METALLOPROTEINASE AND ENCODING DNA THEREFOR  
; FILE REFERENCE: 55-290P

; CURRENT APPLICATION NUMBER: US/09/689,730  
 ; CURRENT FILING DATE: 2000-10-13  
 ; PRIOR APPLICATION NUMBER: US/08/448,489  
 ; PRIOR FILING DATE: 1995-06-07  
 ; NUMBER OF SEQ ID NOS: 19  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 1  
 ; LENGTH: 582  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-09-689-730-1

Query Match 53.9%; Score 41; DB 4; Length 582;  
 Best Local Similarity 81.8%; Pred. No. 1.1e+02;  
 Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 6 AAAAVALPVL 16  
 :|||:||||  
 Db 538 SAAAVLPVL 548

RESULT 49  
 US-09-270-767-41668  
 ; Sequence 41668, Application US/09270767  
 ; Patent No. 6703491  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Homburger et al.  
 ; TITLE OF INVENTION: Nucleic acids and proteins of *Drosophila melanogaster*  
 ; FILE REFERENCE: File Reference: 7326-094  
 ; CURRENT APPLICATION NUMBER: US/09/270,767  
 ; CURRENT FILING DATE: 1999-03-17  
 ; NUMBER OF SEQ ID NOS: 62517  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 41668  
 ; LENGTH: 716  
 ; TYPE: PRT  
 ; ORGANISM: *Drosophila melanogaster*  
 ; FEATURE:  
 ; OTHER INFORMATION: Xaa means any amino acid  
 US-09-270-767-41668

Query Match 53.3%; Score 40.5; DB 4; Length 716;  
 Best Local Similarity 42.9%; Pred. No. 1.6e+02;  
 Matches 12; Conservative 2; Mismatches 1; Indels 13; Gaps 1;

Qy 2 KKRKAA-----AAVLPVL 16  
 ||:||||  
 Db 231 KKKKAKILGPILALVALKAAALPLLL 258

RESULT 50  
 US-09-252-991A-20037  
 ; Sequence 20037, Application US/09252991A  
 ; Patent No. 6551795  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Marc J. Rubenfield et al.  
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
 ; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
 ; FILE REFERENCE: 107196.136  
 ; CURRENT APPLICATION NUMBER: US/09/252,991A  
 ; CURRENT FILING DATE: 1999-02-18  
 ; PRIOR APPLICATION NUMBER: US 60/074,788  
 ; PRIOR FILING DATE: 1998-02-18  
 ; PRIOR APPLICATION NUMBER: US 60/094,190  
 ; PRIOR FILING DATE: 1998-07-27  
 ; NUMBER OF SEQ ID NOS: 33142  
 ; SEQ ID NO 20037  
 ; LENGTH: 238  
 ; TYPE: PRT  
 ; ORGANISM: *Pseudomonas aeruginosa*  
 US-09-252-991A-20037

Query Match 52.6%; Score 40; DB 4; Length 238;

Best Local Similarity 50.0%; Pred. No. 63;  
 Matches 8; Conservative 6; Mismatches 2; Indels 0; Gaps 0;  
 Qy 1 KKKRKAATAAVLPVL 16  
 :||:||||:|  
 Db 45 EEERLAAAAAVPVL 60

Search completed: June 2, 2005, 01:40:43  
 Job time : 36.875 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 2, 2005, 01:31:46 ; Search time 99.3438 Seconds  
(without alignments)  
59.154 Million cell updates/sec

Title: US-10-764-235-24

Perfect score: 76  
Sequence: 1 KKRRKAAAVLLPVLIA 17

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 145611 seqs, 345679903 residues

Total number of hits satisfying chosen parameters: 1465611

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 750 summaries

Database : Published Applications AA:\*

1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/PTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep.\*  
17: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep.\*  
19: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
20: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	76	100.0	17	14	US-10-126-845-24
2	76	100.0	17	14	US-10-126-845-82
3	76	100.0	17	15	US-10-116-275-112
4	76	100.0	17	16	US-10-764-235-24
5	76	100.0	17	17	US-10-955-656-24
6	76	100.0	17	17	US-10-955-656-82
7	58	76.3	17	15	US-10-136-187-30
8	58	76.3	17	17	US-10-850-873-30
9	56	73.7	19	14	US-10-126-845-108
10	56	73.7	19	17	US-10-955-656-108
11	55	72.4	20	14	US-10-126-845-106
12	55	72.4	20	17	US-10-955-656-106
13	53.5	70.4	29	15	US-10-126-103-124
					Sequence 24, Appl
					Sequence 82, Appl
					Sequence 112, Appl
					Sequence 24, Appl
					Sequence 24, Appl
					Sequence 82, Appl
					Sequence 30, Appl
					Sequence 30, Appl
					Sequence 108, Appl
					Sequence 108, Appl
					Sequence 106, Appl
					Sequence 106, Appl
					Sequence 124, Appl

14	53.5	70.4	29	15	US-10-431-096-124	Sequence 124, Appl
15	53.5	70.4	29	17	US-10-858-367-60	Sequence 60, Appl
16	53	69.7	26	14	US-10-126-845-48	Sequence 48, Appl
17	53	69.7	16	15	US-10-136-187-3	Sequence 3, Appl
18	53	69.7	16	15	US-10-136-187-7	Sequence 7, Appl
19	53	69.7	16	15	US-10-136-187-28	Sequence 28, Appl
20	53	69.7	16	15	US-10-116-275-91	Sequence 91, Appl
21	53	69.7	16	16	US-10-764-235-3	Sequence 3, Appl
22	53	69.7	16	16	US-10-764-235-48	Sequence 48, Appl
23	53	69.7	16	17	US-10-850-873-3	Sequence 3, Appl
24	53	69.7	16	17	US-10-850-873-7	Sequence 7, Appl
25	53	69.7	16	17	US-10-850-873-28	Sequence 28, Appl
26	53	69.7	16	17	US-10-955-656-48	Sequence 48, Appl
27	53	69.7	20	15	US-10-136-187-5	Sequence 5, Appl
28	53	69.7	20	15	US-10-136-187-9	Sequence 9, Appl
29	53	69.7	20	17	US-10-850-873-5	Sequence 5, Appl
30	53	69.7	20	17	US-10-850-873-9	Sequence 9, Appl
31	53	69.7	21	14	US-10-126-845-109	Sequence 109, Appl
32	53	69.7	21	14	US-10-126-845-111	Sequence 111, Appl
33	53	69.7	21	17	US-10-955-656-109	Sequence 109, Appl
34	53	69.7	21	17	US-10-955-656-111	Sequence 111, Appl
35	51	67.1	15	14	US-10-126-845-2	Sequence 2, Appl
36	51	67.1	15	14	US-10-126-845-3	Sequence 3, Appl
37	51	67.1	15	14	US-10-126-845-60	Sequence 60, Appl
38	51	67.1	15	14	US-10-126-845-61	Sequence 61, Appl
39	51	67.1	15	15	US-10-136-187-6	Sequence 6, Appl
40	51	67.1	15	15	US-10-136-187-6	Sequence 6, Appl
41	51	67.1	15	15	US-10-136-187-12	Sequence 12, Appl
42	51	67.1	15	15	US-10-136-187-13	Sequence 13, Appl
43	51	67.1	15	15	US-10-136-187-31	Sequence 31, Appl
44	51	67.1	15	15	US-10-136-187-33	Sequence 33, Appl
45	51	67.1	15	15	US-10-136-187-34	Sequence 34, Appl
46	51	67.1	15	15	US-10-136-187-35	Sequence 35, Appl
47	51	67.1	15	15	US-10-136-187-36	Sequence 36, Appl
48	51	67.1	15	15	US-10-136-187-37	Sequence 37, Appl
49	51	67.1	15	15	US-10-136-187-38	Sequence 38, Appl
50	51	67.1	15	15	US-10-136-187-39	Sequence 39, Appl
51	51	67.1	15	15	US-10-136-187-40	Sequence 40, Appl
52	51	67.1	15	16	US-10-764-235-2	Sequence 2, Appl
53	51	67.1	15	17	US-10-850-873-2	Sequence 2, Appl
54	51	67.1	15	17	US-10-850-873-6	Sequence 6, Appl
55	51	67.1	15	17	US-10-850-873-12	Sequence 12, Appl
56	51	67.1	15	17	US-10-850-873-13	Sequence 13, Appl
57	51	67.1	15	17	US-10-850-873-31	Sequence 31, Appl
58	51	67.1	15	17	US-10-850-873-33	Sequence 33, Appl
59	51	67.1	15	17	US-10-850-873-34	Sequence 34, Appl
60	51	67.1	15	17	US-10-850-873-35	Sequence 35, Appl
61	51	67.1	15	17	US-10-850-873-36	Sequence 36, Appl
62	51	67.1	15	17	US-10-850-873-37	Sequence 37, Appl
63	51	67.1	15	17	US-10-850-873-38	Sequence 38, Appl
64	51	67.1	15	17	US-10-850-873-39	Sequence 39, Appl
65	51	67.1	15	17	US-10-850-873-40	Sequence 40, Appl
66	51	67.1	15	17	US-10-955-656-2	Sequence 2, Appl
67	51	67.1	15	17	US-10-955-656-3	Sequence 3, Appl
68	51	67.1	15	17	US-10-955-656-60	Sequence 60, Appl
69	51	67.1	15	17	US-10-955-656-61	Sequence 61, Appl
70	51	67.1	16	15	US-10-136-187-32	Sequence 32, Appl
71	51	67.1	16	17	US-10-850-873-32	Sequence 32, Appl
72	51	67.1	19	14	US-10-126-845-4	Sequence 4, Appl
73	51	67.1	19	14	US-10-126-845-62	Sequence 62, Appl
74	51	67.1	19	15	US-10-116-275-92	Sequence 92, Appl
75	51	67.1	19	16	US-10-764-235-4	Sequence 4, Appl
76	51	67.1	19	17	US-10-955-656-4	Sequence 4, Appl
77	51	67.1	19	17	US-10-955-656-62	Sequence 62, Appl
78	51	67.1	20	14	US-10-126-845-107	Sequence 107, Appl
79	51	67.1	20	17	US-10-955-656-107	Sequence 107, Appl
80	51	67.1	21	15	US-10-116-275-90	Sequence 90, Appl
81	49	64.5	20	14	US-10-126-845-110	Sequence 110, Appl
82	49	64.5	20	17	US-10-955-656-110	Sequence 110, Appl
83	47	61.8	14	14	US-10-126-845-7	Sequence 7, Appl
84	47	61.8	14	14	US-10-126-845-65	Sequence 65, Appl
85	47	61.8	14	15	US-10-116-275-95	Sequence 95, Appl
86	47	61.8	14	16	US-10-764-235-7	Sequence 7, Appl

87	47	61.8	14	16	US-10-764-235-58	Sequence 58, Appl	160	43	56.6	13	16	US-10-764-235-6	Sequence 6, Appl
88	47	61.8	14	17	US-10-955-656-7	Sequence 7, Appl	161	43	56.6	13	17	US-10-955-656-6	Sequence 6, Appl
89	47	61.8	14	17	US-10-955-656-65	Sequence 65, Appl	162	43	56.6	13	17	US-10-955-656-64	Sequence 64, Appl
90	47	61.8	16	14	US-10-126-845-5	Sequence 5, Appl	163	43	56.6	17	14	US-10-126-845-23	Sequence 23, Appl
91	47	61.8	16	14	US-10-126-845-63	Sequence 63, Appl	164	43	56.6	17	14	US-10-126-845-81	Sequence 81, Appl
92	47	61.8	16	15	US-10-116-275-93	Sequence 93, Appl	165	43	56.6	17	15	US-10-116-275-111	Sequence 111, Appl
93	47	61.8	16	16	US-10-764-235-5	Sequence 5, Appl	166	43	56.6	17	16	US-10-764-235-23	Sequence 23, Appl
94	47	61.8	16	17	US-10-955-656-5	Sequence 5, Appl	167	43	56.6	17	16	US-10-652-864-2	Sequence 2, Appl
95	47	61.8	16	17	US-10-955-656-63	Sequence 63, Appl	168	43	56.6	17	17	US-10-795-081A-8	Sequence 8, Appl
96	47	61.8	16	16	US-10-784-309A-1	Sequence 1, Appl	169	43	56.6	17	17	US-10-955-656-23	Sequence 23, Appl
97	47	61.8	18	16	US-10-784-309A-2	Sequence 2, Appl	170	43	56.6	17	17	US-10-955-656-81	Sequence 81, Appl
98	46	60.5	15	15	US-10-136-187-4	Sequence 4, Appl	171	43	56.6	386	10	US-09-948-193-1	Sequence 1, Appl
99	46	60.5	15	15	US-10-136-187-8	Sequence 8, Appl	172	42	55.3	102	16	US-10-437-963-110047	Sequence 110047,
100	46	60.5	15	15	US-10-136-187-29	Sequence 29, Appl	173	42	55.3	295	15	US-10-767-701-38924	Sequence 38924, A
101	46	60.5	15	17	US-10-850-873-4	Sequence 4, Appl	174	42	55.3	295	15	US-10-425-114-60696	Sequence 60696, A
102	46	60.5	15	17	US-10-850-873-8	Sequence 8, Appl	175	42	55.3	470	16	US-10-437-963-197103	Sequence 197103,
103	46	60.5	15	17	US-10-850-873-29	Sequence 29, Appl	176	42	55.3	364	16	US-10-437-963-136955	Sequence 136955,
104	46	60.5	365	15	US-10-282-122A-47524	Sequence 47524, A	177	41	53.9	372	16	US-10-276-272A-9	Sequence 9, Appl
105	44	57.9	17	14	US-10-126-845-104	Sequence 104, App	178	41	53.9	372	16	US-10-276-272A-3	Sequence 3, Appl
106	44	57.9	17	17	US-10-955-656-104	Sequence 104, App	179	41	53.9	381	16	US-10-276-272A-10	Sequence 10, Appl
107	44	57.9	802	15	US-10-259-194A-216	Sequence 216, App	180	41	53.9	389	16	US-10-276-272A-4	Sequence 4, Appl
108	44	57.9	802	16	US-10-437-963-115028	Sequence 115028,	181	41	53.9	391	16	US-10-276-272A-5	Sequence 5, Appl
109	43	56.6	10	13	US-10-116-288-4	Sequence 4, Appl	182	41	53.9	528	9	US-09-884-752A-190	Sequence 190, App
110	43	56.6	10	17	US-10-634-645-4	Sequence 4, Appl	183	41	53.9	528	15	US-10-282-122A-58715	Sequence 58715, A
111	43	56.6	11	13	US-10-116-288-5	Sequence 5, Appl	184	41	53.9	582	9	US-09-916-658-4	Sequence 4, Appl
112	43	56.6	11	14	US-10-126-845-15	Sequence 15, Appl	185	41	53.9	582	9	US-09-391-104-28	Sequence 28, Appl
113	43	56.6	11	14	US-10-126-845-73	Sequence 73, Appl	186	41	53.9	582	9	US-09-801-196-27	Sequence 27, Appl
114	43	56.6	11	15	US-10-116-275-103	Sequence 103, App	187	41	53.9	582	9	US-09-919-497-84	Sequence 84, Appl
115	43	56.6	11	16	US-10-764-235-15	Sequence 15, Appl	188	41	53.9	582	10	US-09-916-849A-2	Sequence 2, Appl
116	43	56.6	11	17	US-10-634-645-5	Sequence 5, Appl	189	41	53.9	582	14	US-10-133-797-6	Sequence 6, Appl
117	43	56.6	11	17	US-10-955-656-15	Sequence 15, Appl	190	41	53.9	582	14	US-10-131-985-43	Sequence 43, Appl
118	43	56.6	11	17	US-10-955-656-73	Sequence 73, Appl	191	41	53.9	582	15	US-10-411-010-26	Sequence 26, Appl
119	43	56.6	12	9	US-09-789-836-12	Sequence 12, Appl	192	41	53.9	582	15	US-10-447-315-9	Sequence 9, Appl
120	43	56.6	12	9	US-09-785-802A-10	Sequence 10, Appl	193	41	53.9	582	16	US-10-276-272A-19	Sequence 19, Appl
121	43	56.6	12	10	US-09-997-465B-4	Sequence 4, Appl	194	41	53.9	582	17	US-10-901-417-43	Sequence 43, Appl
122	43	56.6	12	13	US-10-116-288-1	Sequence 1, Appl	195	41	53.9	584	16	US-10-953-264-26	Sequence 26, Appl
123	43	56.6	12	14	US-10-226-956-286	Sequence 286, App	196	41	53.9	584	16	US-10-276-272A-11	Sequence 11, Appl
124	43	56.6	12	14	US-10-077-555-1	Sequence 1, Appl	197	41	53.9	590	16	US-10-276-272A-12	Sequence 2, Appl
125	43	56.6	12	14	US-10-211-088-304	Sequence 304, App	198	41	53.9	592	16	US-10-276-272A-2	Sequence 12, Appl
126	43	56.6	12	14	US-10-156-570A-27	Sequence 27, Appl	199	41	53.9	968	16	US-10-437-963-134613	Sequence 134613,
127	43	56.6	12	14	US-10-126-845-1	Sequence 1, Appl	200	40	52.6	209	15	US-10-425-114-58425	Sequence 58425, A
128	43	56.6	12	14	US-10-126-845-14	Sequence 14, Appl	201	40	52.6	350	15	US-10-425-114-63670	Sequence 63670, A
129	43	56.6	12	14	US-10-126-845-15	Sequence 15, Appl	202	40	52.6	370	15	US-10-425-114-70773	Sequence 70773, A
130	43	56.6	12	14	US-10-126-845-72	Sequence 72, Appl	203	40	52.6	386	16	US-10-437-963-184361	Sequence 184361,
131	43	56.6	12	14	US-10-126-845-74	Sequence 74, Appl	204	40	52.6	529	15	US-10-282-122A-70592	Sequence 70592, A
132	43	56.6	12	15	US-10-136-187-1	Sequence 1, Appl	205	40	52.6	534	15	US-10-282-122A-70409	Sequence 70409, A
133	43	56.6	12	15	US-10-116-275-102	Sequence 102, App	206	40	52.6	999	16	US-10-437-963-117353	Sequence 117353,
134	43	56.6	12	15	US-10-116-275-104	Sequence 104, App	207	39	51.3	9	13	US-10-116-288-3	Sequence 3, Appl
135	43	56.6	12	15	US-10-144-549-19	Sequence 19, Appl	208	39	51.3	9	14	US-10-126-845-17	Sequence 17, Appl
136	43	56.6	12	15	US-10-361-208-473	Sequence 473, App	209	39	51.3	9	14	US-10-126-845-75	Sequence 75, Appl
137	43	56.6	12	15	US-10-232-410-2	Sequence 2, Appl	210	39	51.3	9	15	US-10-116-275-105	Sequence 105, App
138	43	56.6	12	15	US-10-416-285-1	Sequence 1, Appl	211	39	51.3	9	16	US-10-764-235-17	Sequence 17, Appl
139	43	56.6	12	16	US-10-764-235-1	Sequence 1, Appl	212	39	51.3	9	17	US-10-634-645-3	Sequence 3, Appl
140	43	56.6	12	16	US-10-764-235-14	Sequence 14, Appl	213	39	51.3	9	17	US-10-955-656-17	Sequence 17, Appl
141	43	56.6	12	16	US-10-764-235-16	Sequence 16, Appl	214	39	51.3	9	17	US-10-955-656-75	Sequence 75, Appl
142	43	56.6	12	16	US-10-764-235-59	Sequence 59, Appl	215	39	51.3	11	13	US-10-116-288-9	Sequence 9, Appl
143	43	56.6	12	16	US-10-751-380-8	Sequence 8, Appl	216	39	51.3	11	13	US-10-116-288-9	Sequence 8, Appl
144	43	56.6	12	17	US-10-706-738-1	Sequence 1, Appl	217	39	51.3	11	14	US-10-126-845-19	Sequence 19, Appl
145	43	56.6	12	17	US-10-850-873-1	Sequence 1, Appl	218	39	51.3	11	14	US-10-126-845-66	Sequence 66, Appl
146	43	56.6	12	17	US-10-634-645-1	Sequence 1, Appl	219	39	51.3	11	14	US-10-126-845-77	Sequence 77, Appl
147	43	56.6	12	17	US-10-795-081A-9	Sequence 9, Appl	220	39	51.3	11	15	US-10-116-275-96	Sequence 96, Appl
148	43	56.6	12	17	US-10-823-259-38	Sequence 38, Appl	221	39	51.3	11	15	US-10-116-275-107	Sequence 107, App
149	43	56.6	12	17	US-10-823-254-38	Sequence 38, Appl	222	39	51.3	11	16	US-10-764-235-8	Sequence 8, Appl
150	43	56.6	12	17	US-10-902-959-51	Sequence 51, Appl	223	39	51.3	11	16	US-10-764-235-19	Sequence 19, Appl
151	43	56.6	12	17	US-10-955-656-1	Sequence 1, Appl	224	39	51.3	11	17	US-10-634-645-9	Sequence 9, Appl
152	43	56.6	12	17	US-10-955-656-14	Sequence 14, Appl	225	39	51.3	11	17	US-10-955-656-8	Sequence 8, Appl
153	43	56.6	12	17	US-10-955-656-16	Sequence 16, Appl	226	39	51.3	11	17	US-10-955-656-19	Sequence 19, Appl
154	43	56.6	12	17	US-10-955-656-72	Sequence 72, Appl	227	39	51.3	11	17	US-10-955-656-66	Sequence 66, Appl
155	43	56.6	12	17	US-10-955-656-74	Sequence 74, Appl	228	39	51.3	11	17	US-10-955-656-77	Sequence 77, Appl
156	43	56.6	13	14	US-10-013-845-20	Sequence 20, Appl	229	39	51.3	13	14	US-10-126-845-10	Sequence 10, Appl
157	43	56.6	13	14	US-10-126-845-6	Sequence 6, Appl	230	39	51.3	13	14	US-10-126-845-68	Sequence 68, Appl
158	43	56.6	13	14	US-10-126-845-64	Sequence 64, Appl	231	39	51.3	13	15	US-10-116-275-98	Sequence 98, Appl
159	43	56.6	13	15	US-10-116-275-94	Sequence 94, Appl	232	39	51.3	13	16	US-10-764-235-10	Sequence 10, Appl

233	39	51.3	13	17	US-10-955-656-10	Sequence 10, Appl	306	38	50.0	2802	15	US-10-072-012-489	Sequence 489, App
234	39	51.3	13	17	US-10-955-656-68	Sequence 68, Appl	307	37	48.7	38	10	US-09-892-877-344	Sequence 344, App
235	39	51.3	63	15	US-10-424-599-155550	Sequence 155550,	308	37	48.7	38	10	US-09-948-783-284	Sequence 284, App
236	39	51.3	72	16	US-10-437-963-175664	Sequence 175664,	309	37	48.7	57	16	US-10-437-963-155447	Sequence 155447,
237	39	51.3	119	16	US-10-437-963-191466	Sequence 191466,	310	37	48.7	61	16	US-10-437-963-133086	Sequence 133086,
238	39	51.3	161	15	US-10-264-049-2983	Sequence 2983, Ap	311	37	48.7	71	9	US-09-764-847-981	Sequence 981, App
239	39	51.3	170	16	US-10-767-701-60994	Sequence 60994, A	312	37	48.7	71	14	US-10-092-154-981	Sequence 981, App
240	39	51.3	198	16	US-10-437-963-139840	Sequence 139840,	313	37	48.7	74	15	US-10-424-599-247339	Sequence 247339,
241	39	51.3	230	15	US-10-425-114-70753	Sequence 70753, A	314	37	48.7	74	16	US-10-437-963-168365	Sequence 168365,
242	39	51.3	281	15	US-10-425-114-68479	Sequence 68479, A	315	37	48.7	82	14	US-10-178-213-26	Sequence 26, Appl
243	39	51.3	304	16	US-10-437-963-133811	Sequence 133811,	316	37	48.7	82	14	US-10-178-213-41	Sequence 41, Appl
244	39	51.3	316	16	US-10-437-963-189228	Sequence 189228,	317	37	48.7	90	15	US-10-335-977-4931	Sequence 4931, Ap
245	39	51.3	335	17	US-10-732-923-19017	Sequence 19017, A	318	37	48.7	91	15	US-10-335-977-4930	Sequence 4930, Ap
246	39	51.3	347	15	US-10-425-114-49173	Sequence 49173, A	319	37	48.7	117	16	US-10-437-963-203607	Sequence 203607,
247	39	51.3	385	15	US-10-425-114-56509	Sequence 56509, A	320	37	48.7	136	15	US-10-425-114-47492	Sequence 47492, A
248	39	51.3	1018	16	US-10-437-963-146522	Sequence 146522,	321	37	48.7	137	15	US-10-425-114-41254	Sequence 41254, A
249	39	51.3	1211	14	US-10-156-761-9737	Sequence 9737, Ap	322	37	48.7	154	16	US-10-437-963-121537	Sequence 121537, A
250	39	51.3	1395	17	US-10-732-923-2913	Sequence 2913, Ap	323	37	48.7	167	16	US-10-767-701-46297	Sequence 46297, A
251	39	51.3	1395	17	US-10-732-923-2916	Sequence 2916, Ap	324	37	48.7	177	16	US-10-437-963-148311	Sequence 148311,
252	39	51.3	8917	17	US-10-732-923-20626	Sequence 20626, A	325	37	48.7	183	16	US-10-767-701-46296	Sequence 46296, A
253	38	50.0	20	14	US-10-192-832-54	Sequence 54, Appl	326	37	48.7	186	15	US-10-282-122A-69222	Sequence 69222, A
254	38	50.0	30	14	US-10-192-832-68	Sequence 68, Appl	327	37	48.7	192	15	US-10-425-114-61877	Sequence 61877, A
255	38	50.0	30	14	US-10-192-832-69	Sequence 69, Appl	328	37	48.7	201	16	US-10-437-963-190447	Sequence 190447,
256	38	50.0	30	14	US-10-192-832-72	Sequence 72, Appl	329	37	48.7	202	15	US-10-424-599-177859	Sequence 177859,
257	38	50.0	30	14	US-10-192-832-73	Sequence 73, Appl	330	37	48.7	216	16	US-10-437-963-169924	Sequence 169924,
258	38	50.0	30	15	US-10-467-930-5	Sequence 5, Appli	331	37	48.7	219	16	US-10-437-963-162225	Sequence 162225,
259	38	50.0	51	15	US-10-424-599-151303	Sequence 151303,	332	37	48.7	243	16	US-10-437-963-129795	Sequence 129795,
260	38	50.0	71	9	US-09-815-242-11797	Sequence 11797, A	333	37	48.7	248	14	US-10-146-473-62	Sequence 62, Appl
261	38	50.0	71	14	US-10-127-032-131	Sequence 131, App	334	37	48.7	270	10	US-09-892-877-341	Sequence 341, App
262	38	50.0	71	15	US-10-282-122A-66199	Sequence 66199, A	335	37	48.7	270	10	US-09-948-783-281	Sequence 281, App
263	38	50.0	71	15	US-10-282-122A-69474	Sequence 69474, A	336	37	48.7	275	16	US-10-767-701-43676	Sequence 43676, A
264	38	50.0	78	16	US-10-437-963-133648	Sequence 133648,	337	37	48.7	281	15	US-10-365-742-84	Sequence 84, Appl
265	38	50.0	82	14	US-10-178-213-32	Sequence 32, Appl	338	37	48.7	281	16	US-10-437-963-103020	Sequence 103020,
266	38	50.0	133	14	US-10-029-386-32218	Sequence 32218, A	339	37	48.7	286	9	US-09-712-363-168	Sequence 168, App
267	38	50.0	159	15	US-10-296-115-797	Sequence 797, App	340	37	48.7	288	15	US-10-425-114-41181	Sequence 41181, A
268	38	50.0	164	15	US-10-425-114-37884	Sequence 37884, A	341	37	48.7	302	16	US-10-437-963-185224	Sequence 185224,
269	38	50.0	172	16	US-10-437-963-195077	Sequence 195077,	342	37	48.7	319	15	US-10-425-114-70730	Sequence 70730, A
270	38	50.0	173	15	US-10-424-599-174070	Sequence 174070,	343	37	48.7	319	15	US-10-425-114-70731	Sequence 70731, A
271	38	50.0	195	16	US-10-437-963-166828	Sequence 166828,	344	37	48.7	319	15	US-10-425-114-70732	Sequence 70732, A
272	38	50.0	198	16	US-10-437-963-190094	Sequence 190094,	345	37	48.7	320	15	US-10-425-114-56947	Sequence 56947, A
273	38	50.0	207	16	US-10-767-701-44859	Sequence 44859, A	346	37	48.7	323	15	US-10-425-114-39457	Sequence 39457, A
274	38	50.0	208	15	US-10-425-114-61271	Sequence 61271, A	347	37	48.7	326	16	US-10-437-963-187168	Sequence 187168,
275	38	50.0	209	16	US-10-767-701-56995	Sequence 56995, A	348	37	48.7	343	15	US-10-425-114-65137	Sequence 65137, A
276	38	50.0	228	16	US-10-437-963-186822	Sequence 186822,	349	37	48.7	344	15	US-10-282-122A-61728	Sequence 61728, A
277	38	50.0	302	15	US-10-425-114-48603	Sequence 48603, A	350	37	48.7	350	15	US-10-437-963-204222	Sequence 204222,
278	38	50.0	305	15	US-10-282-122A-63318	Sequence 63318, A	351	37	48.7	353	15	US-10-425-114-46361	Sequence 46361, A
279	38	50.0	307	16	US-10-437-963-104711	Sequence 104711,	352	37	48.7	357	15	US-10-104-047-3150	Sequence 3150, Ap
280	38	50.0	309	14	US-10-304-928-10	Sequence 10, Appl	353	37	48.7	379	16	US-10-437-963-170648	Sequence 170648,
281	38	50.0	337	16	US-10-437-963-113618	Sequence 113618,	354	37	48.7	431	16	US-10-437-963-114382	Sequence 114382,
282	38	50.0	356	15	US-10-437-963-115954	Sequence 115954,	355	37	48.7	455	16	US-10-437-963-202044	Sequence 202044,
283	38	50.0	357	17	US-10-376-397A-44	Sequence 44, Appl	356	37	48.7	466	15	US-10-425-114-37460	Sequence 37460, A
284	38	50.0	357	17	US-10-889-527-44	Sequence 44, Appl	357	37	48.7	469	15	US-10-094-749-2536	Sequence 2536, Ap
285	38	50.0	378	15	US-10-282-122A-71969	Sequence 71969, A	358	37	48.7	495	15	US-10-425-114-44627	Sequence 44627, A
286	38	50.0	383	15	US-10-425-114-49307	Sequence 49307, A	359	37	48.7	499	15	US-10-335-977-4932	Sequence 4932, Ap
287	38	50.0	406	16	US-10-767-701-42204	Sequence 42204, A	360	37	48.7	501	15	US-10-287-971-156	Sequence 156, App
288	38	50.0	414	15	US-10-425-114-50726	Sequence 50726, A	361	37	48.7	501	15	US-10-287-971-158	Sequence 158, App
289	38	50.0	418	9	US-09-966-921A-2	Sequence 2, Appli	362	37	48.7	501	15	US-10-287-971-160	Sequence 160, App
290	38	50.0	515	17	US-10-732-923-7582	Sequence 7582, Ap	363	37	48.7	505	15	US-10-287-971-116	Sequence 116, App
291	38	50.0	515	17	US-10-732-923-7583	Sequence 7583, Ap	364	37	48.7	505	15	US-10-287-971-118	Sequence 118, App
292	38	50.0	523	15	US-10-425-114-59354	Sequence 59354, A	365	37	48.7	505	15	US-10-287-971-120	Sequence 120, App
293	38	50.0	526	16	US-10-437-963-190930	Sequence 190930,	366	37	48.7	505	15	US-10-287-971-122	Sequence 122, App
294	38	50.0	538	16	US-10-437-963-113012	Sequence 113012,	367	37	48.7	505	15	US-10-287-971-124	Sequence 124, App
295	38	50.0	548	15	US-10-282-122A-64041	Sequence 64041, A	368	37	48.7	505	15	US-10-287-971-126	Sequence 126, App
296	38	50.0	566	16	US-10-437-963-113009	Sequence 113009,	369	37	48.7	505	15	US-10-287-971-128	Sequence 128, App
297	38	50.0	578	16	US-10-437-963-157831	Sequence 157831,	370	37	48.7	505	15	US-10-287-971-130	Sequence 130, App
298	38	50.0	627	17	US-10-732-923-9530	Sequence 9530, Ap	371	37	48.7	505	15	US-10-287-971-132	Sequence 132, App
299	38	50.0	657	16	US-10-437-963-163549	Sequence 163549,	372	37	48.7	505	15	US-10-287-971-134	Sequence 134, App
300	38	50.0	1122	17	US-10-794-514A-338	Sequence 338, App	373	37	48.7	505	15	US-10-287-971-136	Sequence 136, App
301	38	50.0	1696	16	US-10-408-765A-822	Sequence 822, App	374	37	48.7	505	15	US-10-287-971-152	Sequence 152, App
302	38	50.0	1720	15	US-10-161-927-8	Sequence 8, Appli	375	37	48.7	524	15	US-10-287-971-114	Sequence 114, App
303	38	50.0	2518	17	US-10-732-923-20624	Sequence 20624, A	376	37	48.7	528	15	US-10-335-977-4933	Sequence 4933, Ap
304	38	50.0	2802	9	US-09-808-602-81	Sequence 81, Appl	377	37	48.7	528	17	US-10-662-126-31	Sequence 31, Appl
305	38	50.0	2802	10	US-09-800-198-69	Sequence 69, Appl	378	37	48.7	532	15	US-10-282-122A-71488	Sequence 71488, A

379	37	48.7	550	10	US-09-882-227-62	Sequence 62, Appl	452	267	10	US-09-924-340-80	Sequence 80, Appl
380	37	48.7	554	15	US-10-094-749-2500	Sequence 2500, Ap	453	267	10	US-09-992-095B-80	Sequence 80, Appl
381	37	48.7	561	15	US-10-451-010-3	Sequence 3, Appl	454	267	10	US-09-999-570-80	Sequence 80, Appl
382	37	48.7	561	17	US-10-491-213-50	Sequence 50, Appl	455	267	14	US-10-000-489-80	Sequence 80, Appl
383	37	48.7	562	15	US-10-262-794A-30	Sequence 30, Appl	456	267	14	US-10-000-986-80	Sequence 80, Appl
384	37	48.7	572	15	US-10-424-599-171078	Sequence 171078,	457	267	14	US-10-154-678-80	Sequence 80, Appl
385	37	48.7	583	37	US-10-106-698-6224	Sequence 6224, Ap	458	267	17	US-10-838-854-80	Sequence 80, Appl
386	37	48.7	614	14	US-10-243-475-1119	Sequence 119, App	459	267	15	US-10-838-854-80	Sequence 59496, A
387	37	48.7	626	16	US-10-437-963-165114	Sequence 165114,	460	273	15	US-10-425-114-59495	Sequence 59496, A
388	37	48.7	653	14	US-10-060-830-1114	Sequence 1114, Ap	461	281	15	US-10-425-114-67615	Sequence 67615, A
389	37	48.7	729	14	US-10-060-830-3	Sequence 3, Appl	462	298	15	US-10-425-114-67448	Sequence 67448, A
390	37	48.7	741	15	US-10-181-108-19	Sequence 19, Appl	463	304	9	US-09-893-737-304	Sequence 304, App
391	37	48.7	743	16	US-10-437-963-204128	Sequence 204128,	464	304	9	US-09-808-602-56	Sequence 56, Appl
392	37	48.7	760	10	US-09-948-783-130	Sequence 130, App	465	304	10	US-09-813-432-6	Sequence 6, Appl
393	37	48.7	760	14	US-10-288-232-19	Sequence 19, Appl	466	304	10	US-09-813-432-34	Sequence 34, Appl
394	37	48.7	760	16	US-10-398-038-19	Sequence 19, Appl	467	304	14	US-10-227-884-84	Sequence 84, Appl
395	37	48.7	760	17	US-10-976-440-19	Sequence 19, Appl	468	304	14	US-10-230-163-84	Sequence 84, Appl
396	37	48.7	763	16	US-10-437-963-120637	Sequence 120637,	469	304	14	US-10-230-338-84	Sequence 84, Appl
397	37	48.7	766	15	US-10-425-114-49492	Sequence 49492, A	470	304	14	US-10-218-631-84	Sequence 84, Appl
398	37	48.7	768	15	US-10-264-049-2269	Sequence 2269, Ap	471	304	14	US-10-230-414-84	Sequence 84, Appl
399	37	48.7	769	14	US-10-191-436-5	Sequence 5, Appl	472	304	14	US-10-232-224-84	Sequence 84, Appl
400	37	48.7	769	14	US-10-191-436-8	Sequence 8, Appl	473	304	14	US-10-216-159A-84	Sequence 84, Appl
401	37	48.7	775	14	US-10-191-436-2	Sequence 2, Appl	474	304	14	US-10-218-849-84	Sequence 84, Appl
402	37	48.7	775	16	US-10-648-593-180	Sequence 180, App	475	304	14	US-10-227-883-84	Sequence 84, Appl
403	37	48.7	777	16	US-10-437-963-200061	Sequence 200061,	476	304	14	US-10-219-076-84	Sequence 84, Appl
404	37	48.7	779	15	US-10-424-599-237524	Sequence 237524,	477	304	14	US-10-219-076-84	Sequence 84, Appl
405	37	48.7	1116	15	US-10-320-797-3337	Sequence 3337, Ap	478	304	14	US-10-230-434-84	Sequence 84, Appl
406	37	48.7	1189	15	US-10-262-794A-26	Sequence 26, Appl	479	304	14	US-10-219-003-84	Sequence 84, Appl
407	37	48.7	1722	16	US-10-437-963-202490	Sequence 202490,	480	304	14	US-10-219-075-84	Sequence 84, Appl
408	36	47.4	55	15	US-10-264-049-3233	Sequence 3233, Ap	481	304	14	US-10-219-464-84	Sequence 84, Appl
409	36	47.4	57	14	US-10-080-254-59	Sequence 59, Appl	482	304	14	US-10-219-479-84	Sequence 84, Appl
410	36	47.4	57	15	US-10-242-355-37	Sequence 37, App	483	304	14	US-10-219-479-84	Sequence 84, Appl
411	36	47.4	57	16	US-10-437-963-156090	Sequence 156090,	484	304	14	US-10-219-481-84	Sequence 84, Appl
412	36	47.4	59	16	US-10-437-963-185601	Sequence 185601,	485	304	14	US-10-230-260-84	Sequence 84, Appl
413	36	47.4	62	9	US-09-764-878-130	Sequence 130, App	486	304	14	US-10-232-231-84	Sequence 84, Appl
414	36	47.4	62	14	US-10-079-854-130	Sequence 465, App	487	304	14	US-10-232-233-84	Sequence 84, Appl
415	36	47.4	62	14	US-10-074-095-455	Sequence 130, App	488	304	14	US-10-216-165-84	Sequence 84, Appl
416	36	47.4	62	15	US-10-212-872-465	Sequence 465, App	489	304	14	US-10-218-956-84	Sequence 84, Appl
417	36	47.4	62	15	US-10-437-963-141350	Sequence 465, App	490	304	14	US-10-219-468-84	Sequence 84, Appl
418	36	47.4	69	16	US-10-437-963-148759	Sequence 141350,	491	304	14	US-10-219-536-84	Sequence 84, Appl
419	36	47.4	73	15	US-10-424-599-266719	Sequence 148759,	492	304	14	US-10-219-536-84	Sequence 84, Appl
420	36	47.4	73	15	US-10-424-599-266719	Sequence 266719,	493	304	14	US-10-233-205-84	Sequence 84, Appl
421	36	47.4	75	15	US-10-424-599-194970	Sequence 194970,	494	304	14	US-10-219-072-84	Sequence 84, Appl
422	36	47.4	97	15	US-10-424-599-244070	Sequence 194970,	495	304	14	US-10-219-474-84	Sequence 84, Appl
423	36	47.4	101	15	US-10-311-129-35	Sequence 244070,	496	304	14	US-10-219-474-84	Sequence 84, Appl
424	36	47.4	104	9	US-09-925-299-1089	Sequence 35, Appl	497	304	14	US-10-219-524-84	Sequence 84, Appl
425	36	47.4	104	10	US-09-925-299-1089	Sequence 1089, Ap	498	304	14	US-10-219-524-84	Sequence 84, Appl
426	36	47.4	106	14	US-10-106-698-4774	Sequence 1089, Ap	499	304	14	US-10-227-880-84	Sequence 84, Appl
427	36	47.4	111	16	US-10-437-963-157472	Sequence 4774, Ap	500	304	14	US-10-227-881-84	Sequence 84, Appl
428	36	47.4	111	16	US-10-437-963-157472	Sequence 167472,	501	304	14	US-10-227-882-84	Sequence 84, Appl
429	36	47.4	135	16	US-10-437-963-197945	Sequence 197945,	502	304	14	US-10-230-436-84	Sequence 84, Appl
430	36	47.4	135	16	US-10-767-701-34868	Sequence 34868, A	503	304	14	US-10-232-223-84	Sequence 84, Appl
431	36	47.4	145	16	US-10-109-048-949	Sequence 44554, A	504	304	14	US-10-232-223-84	Sequence 84, Appl
432	36	47.4	148	16	US-10-437-963-172585	Sequence 949, App	505	304	14	US-10-232-227-84	Sequence 84, Appl
433	36	47.4	160	15	US-10-424-599-204627	Sequence 172585,	506	304	14	US-10-232-229-84	Sequence 84, Appl
434	36	47.4	163	16	US-10-767-701-44185	Sequence 204627,	507	304	14	US-10-233-234-84	Sequence 84, Appl
435	36	47.4	177	15	US-10-425-114-61784	Sequence 44185, A	508	304	14	US-10-219-060-84	Sequence 84, Appl
436	36	47.4	195	15	US-10-282-122A-74812	Sequence 61784, A	509	304	14	US-10-219-060-84	Sequence 366, App
437	36	47.4	197	15	US-10-311-129-2	Sequence 2, Appl	510	304	14	US-10-223-085-366	Sequence 84, Appl
438	36	47.4	197	17	US-10-732-923-1819	Sequence 1819, Ap	511	304	14	US-10-216-162-84	Sequence 84, Appl
439	36	47.4	206	15	US-10-374-780A-490	Sequence 490, App	512	304	14	US-10-216-164-84	Sequence 84, Appl
440	36	47.4	222	15	US-10-412-699B-990	Sequence 990, App	513	304	14	US-10-216-168-84	Sequence 84, Appl
441	36	47.4	222	16	US-10-437-963-104665	Sequence 990, App	514	304	14	US-10-219-055-84	Sequence 84, Appl
442	36	47.4	222	16	US-10-437-963-104665	Sequence 104665,	515	304	14	US-10-219-071-84	Sequence 84, Appl
443	36	47.4	233	15	US-10-374-780A-488	Sequence 488, App	516	304	14	US-10-219-074-84	Sequence 84, Appl
444	36	47.4	233	15	US-10-412-699B-988	Sequence 988, App	517	304	14	US-10-219-077-84	Sequence 84, Appl
445	36	47.4	233	16	US-10-437-963-128938	Sequence 128938,	518	304	14	US-10-219-465-84	Sequence 84, Appl
446	36	47.4	233	17	US-10-732-923-5833	Sequence 5833, Ap	519	304	14	US-10-219-465-84	Sequence 84, Appl
447	36	47.4	236	16	US-10-437-963-204477	Sequence 204477,	520	304	14	US-10-219-469-84	Sequence 84, Appl
448	36	47.4	242	16	US-10-437-963-139301	Sequence 139301,	521	304	14	US-10-219-471-84	Sequence 84, Appl
449	36	47.4	260	15	US-10-282-122A-65468	Sequence 65468, A	522	304	14	US-10-219-473-84	Sequence 84, Appl
450	36	47.4	264	16	US-10-437-963-204731	Sequence 204731,	523	304	14	US-10-219-476-84	Sequence 84, Appl
451	36	47.4	267	10	US-09-992-600A-80	Sequence 80, Appl	524	304	14	US-10-219-482-84	Sequence 84, Appl

525	36	47.4	304	14	US-10-227-874-84	Sequence 84, Appl	598	36	47.4	352	17	US-10-937-758A-121	Sequence 121, App
526	36	47.4	304	14	US-10-227-876-84	Sequence 84, Appl	599	36	47.4	353	15	US-10-424-599-148890	Sequence 148890,
527	36	47.4	304	14	US-10-227-878-84	Sequence 84, Appl	600	36	47.4	355	16	US-10-437-963-186113	Sequence 186113,
528	36	47.4	304	14	US-10-229-974-84	Sequence 84, Appl	601	36	47.4	358	15	US-10-369-493-23251	Sequence 23251, A
529	36	47.4	304	14	US-10-230-024-84	Sequence 84, Appl	602	36	47.4	359	15	US-10-425-114-72928	Sequence 72928, A
530	36	47.4	304	14	US-10-230-113-84	Sequence 84, Appl	603	36	47.4	360	15	US-10-425-114-51856	Sequence 51856, A
531	36	47.4	304	14	US-10-230-183-84	Sequence 84, Appl	604	36	47.4	361	16	US-10-437-963-127459	Sequence 127459, A
532	36	47.4	304	14	US-10-230-234-84	Sequence 84, Appl	605	36	47.4	366	15	US-10-282-122A-68759	Sequence 68759, A
533	36	47.4	304	14	US-10-230-306-84	Sequence 84, Appl	606	36	47.4	411	16	US-10-437-963-142252	Sequence 142252,
534	36	47.4	304	14	US-10-230-426-84	Sequence 84, Appl	607	36	47.4	418	16	US-10-156-761-8243	Sequence 8243, Ap
535	36	47.4	304	14	US-10-230-427-84	Sequence 84, Appl	608	36	47.4	421	14	US-10-425-114-43299	Sequence 43299, A
536	36	47.4	304	14	US-10-230-432-84	Sequence 84, Appl	609	36	47.4	434	15	US-10-425-114-56873	Sequence 56873, A
537	36	47.4	304	14	US-10-230-435-84	Sequence 84, Appl	610	36	47.4	440	15	US-10-425-114-71723	Sequence 71723, A
538	36	47.4	304	14	US-10-230-438-84	Sequence 84, Appl	611	36	47.4	440	15	US-10-369-493-17549	Sequence 17549, A
539	36	47.4	304	14	US-10-232-222-84	Sequence 84, Appl	612	36	47.4	445	15	US-10-437-963-136158	Sequence 136158,
540	36	47.4	304	14	US-10-232-222-84	Sequence 84, Appl	613	36	47.4	472	16	US-10-156-761-13178	Sequence 13178, A
541	36	47.4	304	14	US-10-223-084-366	Sequence 366, App	614	36	47.4	475	14	US-10-173-999-44	Sequence 44, Appl
542	36	47.4	304	14	US-10-223-088-366	Sequence 366, App	615	36	47.4	478	15	US-10-399-405-3	Sequence 3, Appli
543	36	47.4	304	14	US-10-223-090-366	Sequence 366, App	616	36	47.4	478	15	US-10-684-206-12	Sequence 12, Appl
544	36	47.4	304	14	US-10-219-070-84	Sequence 84, Appl	617	36	47.4	480	14	US-10-156-761-11800	Sequence 11800, A
545	36	47.4	304	14	US-10-219-472-84	Sequence 84, Appl	618	36	47.4	480	15	US-10-369-493-3486	Sequence 3486, Ap
546	36	47.4	304	14	US-10-219-527-84	Sequence 84, Appl	619	36	47.4	490	15	US-10-369-493-20003	Sequence 20003, A
547	36	47.4	304	14	US-10-227-877-84	Sequence 84, Appl	620	36	47.4	498	15	US-10-282-122A-61917	Sequence 61917, A
548	36	47.4	304	14	US-10-223-087-366	Sequence 366, App	621	36	47.4	503	15	US-10-238-075-614	Sequence 614, App
549	36	47.4	304	14	US-10-223-083-366	Sequence 366, App	622	36	47.4	504	14	US-10-437-963-132188	Sequence 132188,
550	36	47.4	304	14	US-10-216-166-84	Sequence 84, Appl	623	36	47.4	542	16	US-10-437-963-144664	Sequence 144664,
551	36	47.4	304	14	US-10-218-612-84	Sequence 84, Appl	624	36	47.4	567	17	US-10-732-923-1800	Sequence 1800, Ap
552	36	47.4	304	14	US-10-223-089-366	Sequence 366, App	625	36	47.4	578	17	US-10-732-923-1801	Sequence 1801, Ap
553	36	47.4	304	14	US-10-216-163-84	Sequence 84, Appl	626	36	47.4	578	17	US-10-732-923-1812	Sequence 1812, Ap
554	36	47.4	304	14	US-10-418-064-2	Sequence 2, Appli	627	36	47.4	582	17	US-10-437-963-13683	Sequence 13683,
555	36	47.4	304	14	US-10-223-081-366	Sequence 366, App	628	36	47.4	597	16	US-10-732-923-13941	Sequence 13941, A
556	36	47.4	304	14	US-10-218-765-84	Sequence 84, Appl	629	36	47.4	598	17	US-10-156-761-12207	Sequence 12207, A
557	36	47.4	304	14	US-10-219-063-84	Sequence 84, Appl	630	36	47.4	601	14	US-10-732-923-13939	Sequence 13939, A
558	36	47.4	304	14	US-10-219-066-84	Sequence 84, Appl	631	36	47.4	604	17	US-10-425-114-63755	Sequence 63755, A
559	36	47.4	304	14	US-10-219-067-84	Sequence 84, Appl	632	36	47.4	607	15	US-10-425-114-63781	Sequence 63781, A
560	36	47.4	304	14	US-10-219-068-84	Sequence 84, Appl	633	36	47.4	607	15	US-10-425-114-63784	Sequence 63784, A
561	36	47.4	304	14	US-10-219-069-84	Sequence 84, Appl	634	36	47.4	607	15	US-10-424-599-215078	Sequence 215078,
562	36	47.4	304	14	US-10-219-073-84	Sequence 84, Appl	635	36	47.4	623	15	US-10-732-923-13962	Sequence 13962, A
563	36	47.4	304	14	US-10-219-475-84	Sequence 84, Appl	636	36	47.4	625	15	US-10-424-599-208110	Sequence 208110,
564	36	47.4	304	14	US-10-219-480-84	Sequence 84, Appl	637	36	47.4	625	15	US-10-282-122A-71884	Sequence 71884, A
565	36	47.4	304	14	US-10-219-483-84	Sequence 84, Appl	638	36	47.4	695	15	US-10-369-493-18389	Sequence 18389, A
566	36	47.4	304	14	US-10-219-525-84	Sequence 84, Appl	639	36	47.4	735	15	US-10-149-310-40	Sequence 40, Appl
567	36	47.4	304	14	US-10-219-526-84	Sequence 84, Appl	640	36	47.4	802	17	US-10-732-923-13942	Sequence 13942, A
568	36	47.4	304	14	US-10-219-530-84	Sequence 84, Appl	641	36	47.4	811	16	US-10-437-963-181649	Sequence 181649,
569	36	47.4	304	14	US-10-219-531-84	Sequence 84, Appl	642	36	47.4	837	16	US-10-437-963-152789	Sequence 152789,
570	36	47.4	304	14	US-10-219-532-84	Sequence 84, Appl	643	36	47.4	895	16	US-10-437-963-182576	Sequence 182576,
571	36	47.4	304	14	US-10-219-533-84	Sequence 84, Appl	644	36	47.4	977	15	US-10-004-378A-14	Sequence 14, Appl
572	36	47.4	304	14	US-10-230-437-84	Sequence 84, Appl	645	36	47.4	986	17	US-10-336-626-101	Sequence 101, App
573	36	47.4	304	14	US-10-232-228-84	Sequence 84, Appl	646	36	47.4	986	17	US-10-938-061-101	Sequence 101, App
574	36	47.4	304	14	US-10-223-082-366	Sequence 366, App	647	36	47.4	987	15	US-10-395-027-1193	Sequence 1193, Ap
575	36	47.4	304	15	US-10-174-364-6	Sequence 34, Appl	648	36	47.4	987	17	US-10-936-626-100	Sequence 100, App
576	36	47.4	304	15	US-10-305-654-366	Sequence 366, App	649	36	47.4	987	17	US-10-938-061-100	Sequence 100, App
577	36	47.4	304	15	US-10-080-334-289	Sequence 289, App	650	36	47.4	1001	15	US-10-104-047-2283	Sequence 2283, Ap
578	36	47.4	304	15	US-10-232-226-84	Sequence 84, Appl	651	36	47.4	1055	16	US-10-648-593-139	Sequence 139, App
579	36	47.4	304	15	US-10-297-639-1	Sequence 1, Appli	652	36	47.4	1055	16	US-10-757-262-100	Sequence 100, App
580	36	47.4	304	15	US-10-230-130-84	Sequence 84, Appl	653	36	47.4	1310	15	US-10-369-493-5031	Sequence 5031, Ap
581	36	47.4	304	15	US-10-081-056-366	Sequence 366, App	654	36	47.4	1722	15	US-10-373-801-22	Sequence 22, Appl
582	36	47.4	304	15	US-10-219-535-84	Sequence 84, Appl	655	36	47.4	1883	16	US-10-437-963-132501	Sequence 132501,
583	36	47.4	304	15	US-10-232-230-84	Sequence 84, Appl	656	35.5	46.7	22	15	US-10-449-831A-57	Sequence 57, Appl
584	36	47.4	304	15	US-10-246-583-6	Sequence 6, Appli	657	35.5	46.7	23	15	US-10-449-831A-63	Sequence 63, Appl
585	36	47.4	304	15	US-10-246-583-34	Sequence 34, Appl	658	35.5	46.7	830	15	US-10-282-122A-52567	Sequence 52567, A
586	36	47.4	304	15	US-10-311-129-1	Sequence 1, Appli	659	35	46.1	8	13	US-10-116-288-2	Sequence 2, Appli
587	36	47.4	304	15	US-10-119-480-84	Sequence 84, Appl	660	35	46.1	8	14	US-10-126-845-18	Sequence 18, Appl
588	36	47.4	304	16	US-10-689-832-5	Sequence 6, Appli	661	35	46.1	8	15	US-10-126-845-76	Sequence 76, Appl
589	36	47.4	304	16	US-10-689-832-34	Sequence 34, Appl	662	35	46.1	8	15	US-10-116-275-106	Sequence 106, App
590	36	47.4	305	15	US-10-425-114-59211	Sequence 59211, A	663	35	46.1	8	16	US-10-764-235-18	Sequence 2, Appli
591	36	47.4	313	15	US-10-425-114-61764	Sequence 61764, A	664	35	46.1	8	17	US-10-634-645-2	Sequence 2, Appli
592	36	47.4	327	16	US-10-437-963-132627	Sequence 132627,	665	35	46.1	8	17	US-10-955-656-18	Sequence 18, Appl
593	36	47.4	352	9	US-09-870-759-37	Sequence 37, Appl	666	35	46.1	8	17	US-10-955-656-76	Sequence 76, Appl
594	36	47.4	352	9	US-09-870-759-144	Sequence 144, Appl	667	35	46.1	10	13	US-10-116-288-8	Sequence 8, Appli
595	36	47.4	352	10	US-09-751-708A-37	Sequence 37, Appl	668	35	46.1	10	14	US-10-126-845-9	Sequence 9, Appli
596	36	47.4	352	10	US-09-751-708A-144	Sequence 144, App	669	35	46.1	10	14	US-10-126-845-20	Sequence 20, Appl
597	36	47.4	352	17	US-10-937-758A-37	Sequence 37, Appl	670	35	46.1	10	14	US-10-126-845-67	Sequence 67, Appl

671	35	46.1	10	14	US-10-126-845-78	Sequence 78, Appl	744	35	46.1	163	16	US-10-437-963-114067	Sequence 114067,
672	35	46.1	10	15	US-10-116-275-97	Sequence 97, Appl	745	35	46.1	172	15	US-10-108-260A-3030	Sequence 3030, Ap
673	35	46.1	10	15	US-10-116-275-108	Sequence 108, Appl	746	35	46.1	174	15	US-10-425-114-68640	Sequence 68640, A
674	35	46.1	10	16	US-10-764-235-9	Sequence 9, Appl	747	35	46.1	177	15	US-10-424-599-226794	Sequence 226794,
675	35	46.1	10	16	US-10-764-235-20	Sequence 20, Appl	748	35	46.1	184	15	US-10-425-114-39065	Sequence 39065, A
676	35	46.1	10	17	US-10-634-645-8	Sequence 8, Appl	749	35	46.1	187	15	US-10-425-114-44299	Sequence 44299, A
677	35	46.1	10	17	US-10-955-656-9	Sequence 9, Appl	750	35	46.1	193	15	US-10-424-599-152339	Sequence 152339,
678	35	46.1	10	17	US-10-955-656-20	Sequence 20, Appl							
679	35	46.1	10	17	US-10-955-656-67	Sequence 67, Appl							
680	35	46.1	10	17	US-10-955-656-78	Sequence 78, Appl							
681	35	46.1	11	17	US-10-916-185-17	Sequence 17, Appl							
682	35	46.1	12	14	US-10-126-845-11	Sequence 11, Appl							
683	35	46.1	12	14	US-10-126-845-69	Sequence 69, Appl							
684	35	46.1	12	15	US-10-116-275-99	Sequence 99, Appl							
685	35	46.1	12	16	US-10-764-235-11	Sequence 11, Appl							
686	35	46.1	12	17	US-10-955-656-11	Sequence 11, Appl							
687	35	46.1	12	17	US-10-955-656-69	Sequence 69, Appl							
688	35	46.1	51	16	US-10-437-963-123512	Sequence 123512,							
689	35	46.1	52	15	US-10-424-599-173617	Sequence 173617,							
690	35	46.1	54	15	US-10-424-599-234012	Sequence 234012,							
691	35	46.1	58	9	US-09-764-877-1908	Sequence 1908, Ap							
692	35	46.1	58	15	US-10-242-515-1908	Sequence 1908, Ap							
693	35	46.1	58	16	US-10-767-701-47863	Sequence 47863, A							
694	35	46.1	61	16	US-10-437-963-133817	Sequence 133817,							
695	35	46.1	64	15	US-10-424-599-177960	Sequence 177960,							
696	35	46.1	70	15	US-10-424-599-233335	Sequence 233335,							
697	35	46.1	78	15	US-10-424-599-284800	Sequence 284800,							
698	35	46.1	79	15	US-10-276-774-2489	Sequence 2489, Ap							
699	35	46.1	79	16	US-10-437-963-150958	Sequence 150958,							
700	35	46.1	82	16	US-10-437-963-216789	Sequence 216789,							
701	35	46.1	85	15	US-10-424-599-216789	Sequence 216789,							
702	35	46.1	85	16	US-10-437-963-155953	Sequence 155953,							
703	35	46.1	89	16	US-10-437-963-157858	Sequence 157858,							
704	35	46.1	90	16	US-10-767-701-35816	Sequence 35816, A							
705	35	46.1	90	16	US-10-767-701-36332	Sequence 36332, A							
706	35	46.1	91	9	US-09-071-035-382	Sequence 382, App							
707	35	46.1	91	9	US-09-071-035-386	Sequence 386, App							
708	35	46.1	91	14	US-10-206-576-382	Sequence 382, App							
709	35	46.1	91	14	US-10-206-576-386	Sequence 386, App							
710	35	46.1	91	16	US-10-437-963-200604	Sequence 200604,							
711	35	46.1	91	17	US-10-912-362-382	Sequence 382, App							
712	35	46.1	91	17	US-10-912-362-386	Sequence 386, App							
713	35	46.1	100	15	US-10-425-114-47419	Sequence 47419, A							
714	35	46.1	102	15	US-10-424-599-209474	Sequence 209474, A							
715	35	46.1	102	15	US-10-425-114-68452	Sequence 68452, A							
716	35	46.1	106	15	US-10-335-977-5552	Sequence 5552, Ap							
717	35	46.1	107	15	US-10-424-599-226795	Sequence 226795,							
718	35	46.1	107	16	US-10-767-701-56178	Sequence 56178, A							
719	35	46.1	109	15	US-10-424-599-220836	Sequence 220836,							
720	35	46.1	109	16	US-10-437-963-131661	Sequence 131661,							
721	35	46.1	110	17	US-10-732-923-18794	Sequence 18794, A							
722	35	46.1	111	15	US-10-424-599-256435	Sequence 256435,							
723	35	46.1	111	15	US-10-425-114-71124	Sequence 71124, A							
724	35	46.1	113	16	US-10-437-963-182201	Sequence 182201,							
725	35	46.1	117	15	US-10-389-566-377	Sequence 377, App							
726	35	46.1	117	16	US-10-437-963-121323	Sequence 121323,							
727	35	46.1	120	15	US-10-264-049-3501	Sequence 3501, Ap							
728	35	46.1	126	16	US-10-437-963-115152	Sequence 115152,							
729	35	46.1	131	15	US-10-283-122A-60820	Sequence 60820, A							
730	35	46.1	133	16	US-10-767-701-54208	Sequence 54208, A							
731	35	46.1	134	15	US-10-424-599-149365	Sequence 149365,							
732	35	46.1	138	9	US-09-764-869-929	Sequence 929, App							
733	35	46.1	138	14	US-10-091-504-929	Sequence 929, App							
734	35	46.1	138	15	US-10-227-577-929	Sequence 929, App							
735	35	46.1	140	15	US-10-424-599-232686	Sequence 232686,							
736	35	46.1	140	16	US-10-767-701-60606	Sequence 60606, A							
737	35	46.1	142	16	US-10-437-963-136962	Sequence 136962,							
738	35	46.1	142	16	US-10-437-963-149310	Sequence 149310,							
739	35	46.1	144	16	US-10-767-701-36831	Sequence 36831, A							
740	35	46.1	149	16	US-10-767-701-33687	Sequence 33687, A							
741	35	46.1	155	15	US-10-425-114-52725	Sequence 52725, A							
742	35	46.1	162	10	US-09-826-734-112	Sequence 112, App							
743	35	46.1	162	10	US-09-826-734-118	Sequence 118, App							

## ALIGNMENTS

## RESULT 1

US-10-126-845-24  
; Sequence 24, Application US/10126845  
; Publication NO. US20030181367A1  
; GENERAL INFORMATION:  
; APPLICANT: O'Mahony, Daniel J.  
; APPLICANT: Lambkin, Imelda J.  
; APPLICANT: Pinilla, Clemencia  
; APPLICANT: Houghten, Richard  
; TITLE OF INVENTION: MEMBRANE TRANSLOCATING PEPTIDE DRUG DELIVERY SYSTEM  
; FILE REFERENCE: E1067/20058  
; CURRENT FILING DATE: 2002-10-15  
; NUMBER OF SEQ ID NOS: 119  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 24  
; LENGTH: 17  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: membrane translocating peptide  
US-10-126-845-24

Query Match 100.0%; Score 76; DB 14; Length 17;  
Best Local Similarity 100.0%; Pred. No. 4.4e-05;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKKKAAAVLLPVLLA 17  
|||||  
DB 1 KKKKAAAVLLPVLLA 17

## RESULT 2

US-10-126-845-82  
; Sequence 82, Application US/10126845  
; Publication NO. US20030181367A1  
; GENERAL INFORMATION:  
; APPLICANT: O'Mahony, Daniel J.  
; APPLICANT: Lambkin, Imelda J.  
; APPLICANT: Pinilla, Clemencia  
; APPLICANT: Houghten, Richard  
; TITLE OF INVENTION: MEMBRANE TRANSLOCATING PEPTIDE DRUG DELIVERY SYSTEM  
; FILE REFERENCE: E1067/20058  
; CURRENT FILING DATE: 2002-10-15  
; NUMBER OF SEQ ID NOS: 119  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 82  
; LENGTH: 17  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: D form peptide  
; FEATURE:  
; NAME/KEY: MISC FEATURE  
; LOCATION: (1)-(17)  
; OTHER INFORMATION: D form amino acid  
US-10-126-845-82

Query Match 100.0%; Score 76; DB 14; Length 17;  
Best Local Similarity 100.0%; Pred. No. 4.4e-05;



Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKKRKAATAVLLPVLLA 17  
|||||  
Db 1 KKKRKAATAVLLPVLLA 17  
|||||

RESULT 3  
US-10-116-275-112  
; Sequence 112, Application US/10116275  
; Publication No. US20030211476A1  
; GENERAL INFORMATION:  
; APPLICANT: Elan Pharmaceutical Technology  
; APPLICANT: O'Mahony, Daniel J.  
; APPLICANT: Brayden, David  
; APPLICANT: Byrne, Daragh  
; APPLICANT: Lambkin, Imelda  
; APPLICANT: Higgins, Lisa  
; TITLE OF INVENTION: Genetic Analysis of Peyer's Patches and M Cells and Methods and  
; FILE REFERENCE: E1067/20087  
; CURRENT APPLICATION NUMBER: US/10/116,275  
; CURRENT FILING DATE: 2002-10-04  
; NUMBER OF SEQ ID NOS: 349  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 112  
; LENGTH: 17  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Listed in Table titled "Peptides that Target to and/or Enhance Up  
; OTHER INFORMATION: take Across the GIT"  
US-10-116-275-112

Query Match 100.0%; Score 76; DB 15; Length 17;  
Best Local Similarity 100.0%; Pred. No. 4.4e-05;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKKRKAATAVLLPVLLA 17  
|||||  
Db 1 KKKRKAATAVLLPVLLA 17  
|||||

RESULT 4  
US-10-764-235-24  
; Sequence 24, Application US/10764235  
; Publication No. US20040138132A1  
; GENERAL INFORMATION:  
; APPLICANT: O'Mahony, Daniel J.  
; APPLICANT: Lambkin, Imelda J.  
; TITLE OF INVENTION: MEMBRANE TRANSLOCATING PEPTIDE DRUG DELIVERY SYSTEM  
; FILE REFERENCE: P26,479-B USA  
; CURRENT APPLICATION NUMBER: US/10/764,235  
; CURRENT FILING DATE: 2004-01-23  
; PRIOR APPLICATION NUMBER: 09/671,089  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: 60/156,246  
; PRIOR FILING DATE: 1999-09-27  
; NUMBER OF SEQ ID NOS: 59  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 24  
; LENGTH: 17  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: membrane translocating peptide  
US-10-764-235-24

Query Match 100.0%; Score 76; DB 16; Length 17;  
Best Local Similarity 100.0%; Pred. No. 4.4e-05;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKKRKAATAVLLPVLLA 17

Db 1 KKKRKAATAVLLPVLLA 17  
|||||

RESULT 5  
US-10-955-656-24  
; Sequence 24, Application US/10955656  
; Publication No. US20050101762A1  
; GENERAL INFORMATION:  
; APPLICANT: O'Mahony, Daniel J.  
; APPLICANT: Lambkin, Imelda J.  
; APPLICANT: Pinilla, Clemencia  
; APPLICANT: Houghten, Richard  
; TITLE OF INVENTION: CONJUGATES OF MEMBRANE TRANSLOCATING AGENTS AND PHARMACEUTICALLY  
; FILE REFERENCE: P26,481-A USA  
; CURRENT APPLICATION NUMBER: US/10/955,656  
; CURRENT FILING DATE: 2004-09-30  
; PRIOR APPLICATION NUMBER: 10/126,845  
; PRIOR FILING DATE: 2002-04-19  
; PRIOR APPLICATION NUMBER: 09/671,089  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: 60/156,246  
; PRIOR FILING DATE: 1999-09-27  
; NUMBER OF SEQ ID NOS: 119  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 24  
; LENGTH: 17  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: membrane translocating peptide  
US-10-955-656-24

Query Match 100.0%; Score 76; DB 17; Length 17;  
Best Local Similarity 100.0%; Pred. No. 4.4e-05;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KKKRKAATAVLLPVLLA 17  
|||||  
Db 1 KKKRKAATAVLLPVLLA 17  
|||||

RESULT 6  
US-10-955-656-82  
; Sequence 82, Application US/10955656  
; Publication No. US20050101762A1  
; GENERAL INFORMATION:  
; APPLICANT: O'Mahony, Daniel J.  
; APPLICANT: Lambkin, Imelda J.  
; APPLICANT: Pinilla, Clemencia  
; APPLICANT: Houghten, Richard  
; TITLE OF INVENTION: CONJUGATES OF MEMBRANE TRANSLOCATING AGENTS AND PHARMACEUTICALLY  
; FILE REFERENCE: P26,481-A USA  
; CURRENT APPLICATION NUMBER: US/10/955,656  
; CURRENT FILING DATE: 2004-09-30  
; PRIOR APPLICATION NUMBER: 10/126,845  
; PRIOR FILING DATE: 2002-04-19  
; PRIOR APPLICATION NUMBER: 09/671,089  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: 60/156,246  
; PRIOR FILING DATE: 1999-09-27  
; NUMBER OF SEQ ID NOS: 119  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 82  
; LENGTH: 17  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: D form peptide  
; NAME/KEY: MISC\_FEATURE



;; CURRENT FILING DATE: 2004-09-30  
;; PRIOR APPLICATION NUMBER: 10/126,845  
;; PRIOR FILING DATE: 2002-04-19  
;; PRIOR APPLICATION NUMBER: 09/671,089  
;; PRIOR FILING DATE: 2000-09-27  
;; PRIOR APPLICATION NUMBER: 60/156,246  
;; PRIOR FILING DATE: 1999-09-27  
;; NUMBER OF SEQ ID NOS: 119  
;; SOFTWARE: PatentIn version 3.1  
;; SEQ ID NO 108  
;; LENGTH: 19  
;; TYPE: PRT  
;; ORGANISM: Artificial Sequence  
;; OTHER INFORMATION: D-form peptide; Comprises opioid peptide  
;; FEATURE:  
;; NAME/KEY: MISC FEATURE  
;; LOCATION: (1)..(19)  
;; OTHER INFORMATION: D-form amino acid  
;; FEATURE:  
;; NAME/KEY: MOD RES  
;; LOCATION: (19)..(19)  
;; OTHER INFORMATION: Amidation  
;; US-10-955-656-106

Query Match 73.7%; Score 56; DB 17; Length 19;  
Best Local Similarity 92.9%; Pred. No. 0.063;  
Matches 13; Conservative 0; Mismatches 0; Gaps 0;

QY 4 RKAAAALLPVLLA 17  
DB 4 RKAAAALLPVLLA 17

RESULT 11  
US-10-126-845-106  
; Sequence 106, Application US/10126845  
; Publication No. US20030181367A1  
; GENERAL INFORMATION:  
; APPLICANT: O'Mahony, Daniel J.  
; APPLICANT: Lambkin, Imelda J.  
; APPLICANT: Pinilla, Clemencia  
; APPLICANT: Houghten, Richard  
; TITLE OF INVENTION: MEMBRANE TRANSLOCATING PEPTIDE DRUG DELIVERY SYSTEM  
; FILE REFERENCE: E1067/20058  
; CURRENT APPLICATION NUMBER: US/10/126,845  
; CURRENT FILING DATE: 2002-10-15  
; NUMBER OF SEQ ID NOS: 119  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 106  
; LENGTH: 20  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: D-form peptide; Comprises opioid peptide  
; FEATURE:  
; NAME/KEY: MISC FEATURE  
; LOCATION: (1)..(20)  
; OTHER INFORMATION: D-form amino acid  
; FEATURE:  
; NAME/KEY: MOD RES  
; LOCATION: (20)..(20)  
; OTHER INFORMATION: Amidation  
; US-10-126-845-106

Query Match 72.4%; Score 55; DB 14; Length 20;  
Best Local Similarity 80.0%; Pred. No. 0.096;  
Matches 12; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 RKAAAALLPVLLA 17  
DB 4 RKAAAALLPVLLA 18

RESULT 12  
US-10-955-656-106  
; Sequence 106, Application US/10955656  
; Publication No. US20050101762A1  
; GENERAL INFORMATION:  
; APPLICANT: O'Mahony, Daniel J.  
; APPLICANT: Lambkin, Imelda J.  
; APPLICANT: Pinilla, Clemencia  
; APPLICANT: Houghten, Richard  
; TITLE OF INVENTION: CONJUGATES OF MEMBRANE TRANSLOCATING AGENTS AND PHARMACEUTICALLY  
; FILE REFERENCE: P26,481-A USA  
; CURRENT APPLICATION NUMBER: US/10/955,656  
; CURRENT FILING DATE: 2004-09-30  
; PRIOR APPLICATION NUMBER: 10/126,845  
; PRIOR FILING DATE: 2002-04-19  
; PRIOR APPLICATION NUMBER: 09/671,089  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: 60/156,246  
; PRIOR FILING DATE: 1999-09-27  
; NUMBER OF SEQ ID NOS: 119  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 106  
; LENGTH: 20  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: D-form peptide; Comprises opioid peptide  
; FEATURE:  
; NAME/KEY: MISC FEATURE  
; LOCATION: (1)..(20)  
; OTHER INFORMATION: D-form amino acid  
; FEATURE:  
; NAME/KEY: MOD RES  
; LOCATION: (20)..(20)  
; OTHER INFORMATION: Amidation  
; US-10-955-656-106

Query Match 72.4%; Score 55; DB 17; Length 20;  
Best Local Similarity 80.0%; Pred. No. 0.096;  
Matches 12; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 RKAAAALLPVLLA 17  
DB 4 RKAAAALLPVLLA 18

RESULT 13  
US-10-126-103-124  
; Sequence 124, Application US/10126103  
; Publication No. US20030224486A1  
; GENERAL INFORMATION:  
; APPLICANT: Bristol-Myers Squibb Company  
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES ASSOCIATED WITH THE NF-KB PATHWAY  
; FILE REFERENCE: D0108.np  
; CURRENT APPLICATION NUMBER: US/10/126,103  
; CURRENT FILING DATE: 2002-04-19  
; PRIOR APPLICATION NUMBER: US 60/284,962  
; PRIOR FILING DATE: 2001-04-19  
; PRIOR APPLICATION NUMBER: US 60/286,645  
; PRIOR FILING DATE: 2001-04-26  
; PRIOR APPLICATION NUMBER: US 60/346,986  
; PRIOR FILING DATE: 2002-01-09  
; NUMBER OF SEQ ID NOS: 284  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 124  
; LENGTH: 29  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Synthetic peptide sequence.  
; US-10-126-103-124

QY	1	KKRKAATAVLLP-VLLA	17 
Db	2	KKRKVAALLPAVLLA	19 
RESULT 16			
US-10-126--845-48			
; Sequence 48; Application US/10126845			
; Publication No. US20030181367A1			
GENERAL INFORMATION:			
; APPLICANT: O'Mahony, Daniel J.			
; APPLICANT: Lambkin, Inelda J.			
; APPLICANT: Pinilla, Clemencia			
; APPLICANT: Houghten, Richard			
TITLE OF INVENTION: MEMBRANE TRANSLOCATING PEPTIDE DRUG DELIVERY SYSTEM			
FILE REFERENCE: E1067/20058			
CURRENT APPLICATION NUMBER: US/10/126,845			
CURRENT FILING DATE: 2002-10-15			
NUMBER OF SEQ ID NOS: 119			
SOFTWARE: PatentIn version 3.1			
SEQ ID NO 48			
LENGTH: 16			
TYPE: PRT			
ORGANISM: Artificial Sequence			
FEATURE:			
OTHER INFORMATION: dansylated membrane translocating peptide			
FEATURE:			
NAME/KEY: MOD RES			
LOCATION: (1)-(1)			
OTHER INFORMATION: dansylated			
US-10-126-845-48			
Query Match 69.7%; Score 53; DB 14; Length 16;			
Best Local Similarity 85.7%; Pred.No. 0.16;			
Matches 12; Conservative 1; Mismatches 1; Indels 0; Gaps 0;			
QY	4	RKAAAVLPLVLLA	17 :
Db	1	KKKAAVLLPVLIA	14 
RESULT 17			
US-10-136-187-3			
; Sequence 3; Application US/10136187			
; Publication No. US20030203865A1			
GENERAL INFORMATION:			
; APPLICANT: Harvie, Pierrot			
; APPLICANT: Paul, Ralph			
; APPLICANT: Cudmore, Sally			
; APPLICANT: O'Mahony, Daniel J.			
TITLE OF INVENTION: LIPID-COMPRISING DRUG DELIVERY COMPLEXES			
TITLE OF INVENTION: AND METHODS FOR THEIR PRODUCTION			
FILE REFERENCE: 226272005300			
CURRENT APPLICATION NUMBER: US/10/136,187			
CURRENT FILING DATE: 2002-09-13			
PRIOR APPLICATION NUMBER: US 60/287,786			
PRIOR FILING DATE: 2001-04-30			
NUMBER OF SEQ ID NOS: 45			
SOFTWARE: FastSeq for Windows Version 4.0			
SEQ ID NO 3			
LENGTH: 16			
TYPE: PRT			
ORGANISM: Artificial Sequence			
FEATURE:			
OTHER INFORMATION: Synthetic Construct			
US-10-136-187-3			
Query Match 69.7%; Score 53; DB 15; Length 16;			
Best Local Similarity 85.7%; Pred.No. 0.16;			
Matches 12; Conservative 1; Mismatches 1; Indels 0; Gaps 0;			
QY	4	RKAAAVLPLVLLA	17 :

Db 1 KKKAAAVLLPVLLA 14  
:| ||||| |||||

## RESULT 18

US-10-136-187-7  
; Sequence 7, Application US/10136187  
; Publication No. US20030203865A1  
; GENERAL INFORMATION:  
; APPLICANT: Harvie, Pierrot  
; APPLICANT: Paul, Ralph  
; APPLICANT: Cudmore, Sally  
; APPLICANT: O'Mahony, Daniel J.  
; TITLE OF INVENTION: LIPID-COMPRISING DRUG DELIVERY COMPLEXES  
; FILE REFERENCE: 226272005300  
; CURRENT APPLICATION NUMBER: US/10/136,187  
; CURRENT FILING DATE: 2002-09-13  
; PRIOR APPLICATION NUMBER: US 60/287,786  
; PRIOR FILING DATE: 2001-04-30  
; NUMBER OF SEQ ID NOS: 45  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 7  
; LENGTH: 16  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Synthetic Construct  
; NAME/KEY: VARIANT  
; LOCATION: 1  
; OTHER INFORMATION: Lysine is attached to galactose  
US-10-136-187-7

Query Match 69.7%; Score 53; DB 15; Length 16;  
Best Local Similarity 85.7%; Pred. No. 0.16;  
Matches 12; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 KKKAAAVLLPVLLA 17  
:| ||||| |||||

Db 1 KKKAAAVLLPVLLA 14

## RESULT 19

US-10-136-187-28  
; Sequence 28, Application US/10136187  
; Publication No. US20030203865A1  
; GENERAL INFORMATION:  
; APPLICANT: Harvie, Pierrot  
; APPLICANT: Paul, Ralph  
; APPLICANT: Cudmore, Sally  
; APPLICANT: O'Mahony, Daniel J.  
; TITLE OF INVENTION: LIPID-COMPRISING DRUG DELIVERY COMPLEXES  
; FILE REFERENCE: 226272005300  
; CURRENT APPLICATION NUMBER: US/10/136,187  
; CURRENT FILING DATE: 2002-09-13  
; PRIOR APPLICATION NUMBER: US 60/287,786  
; PRIOR FILING DATE: 2001-04-30  
; NUMBER OF SEQ ID NOS: 45  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 28  
; LENGTH: 16  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Synthetic Construct  
; NAME/KEY: VARIANT  
; LOCATION: 1  
; OTHER INFORMATION: Lysine is attached to dansyl  
US-10-136-187-28

Query Match 69.7%; Score 53; DB 15; Length 16;  
Best Local Similarity 85.7%; Pred. No. 0.16;  
Matches 12; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 KKKAAAVLLPVLLA 17  
:| ||||| |||||

Db 1 KKKAAAVLLPVLLA 14

## RESULT 20

US-10-116-275-91  
; Sequence 91, Application US/10116275  
; Publication No. US20030211476A1  
; GENERAL INFORMATION:  
; APPLICANT: Elan Pharmaceutical Technology  
; APPLICANT: O'Mahony, Daniel J.  
; APPLICANT: Brayden, David  
; APPLICANT: Byrne, Daragh  
; APPLICANT: Lambkin, Imelda  
; APPLICANT: Higgins, Lisa  
; TITLE OF INVENTION: Genetic Analysis of Peyer's Patches and M Cells and Methods and  
; FILE REFERENCE: E1067/20087  
; CURRENT APPLICATION NUMBER: US/10/116,275  
; CURRENT FILING DATE: 2002-10-04  
; NUMBER OF SEQ ID NOS: 349  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 91  
; LENGTH: 16  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Listed in Table titled "Peptides that Target to and/or Enhance U;  
; OTHER INFORMATION: take Across the GIT"  
US-10-116-275-91

Query Match 69.7%; Score 53; DB 15; Length 16;  
Best Local Similarity 85.7%; Pred. No. 0.16;  
Matches 12; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 KKKAAAVLLPVLLA 17  
:| ||||| |||||

Db 1 KKKAAAVLLPVLLA 14

## RESULT 21

US-10-764-235-3  
; Sequence 3, Application US/10764235  
; Publication No. US20040138132A1  
; GENERAL INFORMATION:  
; APPLICANT: O'Mahony, Daniel J.  
; APPLICANT: Lambkin, Imelda J.  
; TITLE OF INVENTION: MEMBRANE TRANSLOCATING PEPTIDE DRUG DELIVERY SYSTEM  
; FILE REFERENCE: P26,479-B USA  
; CURRENT APPLICATION NUMBER: US/10/764,235  
; CURRENT FILING DATE: 2004-01-23  
; PRIOR APPLICATION NUMBER: 09/671,089  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: 60/156,246  
; PRIOR FILING DATE: 1999-09-27  
; NUMBER OF SEQ ID NOS: 59  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 3  
; LENGTH: 16  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: membrane translocating peptide  
US-10-764-235-3

Query Match 69.7%; Score 53; DB 16; Length 16;  
Best Local Similarity 85.7%; Pred. No. 0.16;  
Matches 12; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

QY      4 RKAAAALLPVLLA 17
Db      1 KKKAAAALLPVLLA 14

RESULT 22
US-10-764-235-48
; Sequence 48, Application US/10764235
; Publication No. US20040138132A1
; GENERAL INFORMATION:
; APPLICANT: O'Mahony, Daniel J.
; APPLICANT: Lambkin, Inelda J.
; TITLE OF INVENTION: MEMBRANE TRANSLOCATING PEPTIDE DRUG DELIVERY SYSTEM
; FILE REFERENCE: P26,479-B USA
; CURRENT APPLICATION NUMBER: US/10/764,235
; CURRENT FILING DATE: 2004-01-23
; PRIOR APPLICATION NUMBER: 09/671,089
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/156,246
; PRIOR FILING DATE: 1999-09-27
; NUMBER OF SEQ ID NOS: 59
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 48
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: dansylated membrane translocating peptide
; NAME/KEY: MOD RES
; LOCATION: (1)..(1)
; OTHER INFORMATION: dansylated
US-10-764-235-48

Query Match      69.7%; Score 53; DB 16; Length 16;
Best Local Similarity 85.7%; Pred. No. 0.16;
Matches 12; Conservative 1; Mismatches 1; Indels 1; Gaps 0;

QY      4 RKAAAALLPVLLA 17
Db      1 KKKAAAALLPVLLA 14

RESULT 23
US-10-850-873-3
; Sequence 3, Application US/10850873
; Publication No. US20050025821A1
; GENERAL INFORMATION:
; APPLICANT: Harvie, Pierrot
; APPLICANT: Paul, Ralph
; APPLICANT: Cudmore, Sally
; APPLICANT: O'Mahony, Daniel J.
; TITLE OF INVENTION: LIPID-COMPRISING DRUG DELIVERY COMPLEXES
; TITLE OF INVENTION: AND METHODS FOR THEIR PRODUCTION
; FILE REFERENCE: 226272005300
; CURRENT APPLICATION NUMBER: US/10/850,873
; CURRENT FILING DATE: 2004-05-20
; PRIOR APPLICATION NUMBER: US/10/136,187
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/287,786
; PRIOR FILING DATE: 2001-04-30
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Construct
US-10-850-873-3

Query Match      69.7%; Score 53; DB 17; Length 16;
Best Local Similarity 85.7%; Pred. No. 0.16;
Matches 12; Conservative 1; Mismatches 1; Indels 1; Gaps 0;

```

```

Best Local Similarity 85.7%; Pred. No. 0.16;
Matches 12; Conservative 1; Mismatches 1; Indels 1; Gaps 0;

QY      4 RKAAAALLPVLLA 17
Db      1 KKKAAAALLPVLLA 14

RESULT 24
US-10-850-873-7
; Sequence 7, Application US/10850873
; Publication No. US20050025821A1
; GENERAL INFORMATION:
; APPLICANT: Harvie, Pierrot
; APPLICANT: Paul, Ralph
; APPLICANT: Cudmore, Sally
; APPLICANT: O'Mahony, Daniel J.
; TITLE OF INVENTION: LIPID-COMPRISING DRUG DELIVERY COMPLEXES
; TITLE OF INVENTION: AND METHODS FOR THEIR PRODUCTION
; FILE REFERENCE: 226272005300
; CURRENT APPLICATION NUMBER: US/10/850,873
; CURRENT FILING DATE: 2004-05-20
; PRIOR APPLICATION NUMBER: US/10/136,187
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/287,786
; PRIOR FILING DATE: 2001-04-30
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Construct
; NAME/KEY: VARIANT
; LOCATION: 1
; OTHER INFORMATION: Lysine is attached to galactose
US-10-850-873-7

Query Match      69.7%; Score 53; DB 17; Length 16;
Best Local Similarity 85.7%; Pred. No. 0.16;
Matches 12; Conservative 1; Mismatches 1; Indels 1; Gaps 0;

QY      4 RKAAAALLPVLLA 17
Db      1 KKKAAAALLPVLLA 14

RESULT 25
US-10-850-873-28
; Sequence 28, Application US/10850873
; Publication No. US20050025821A1
; GENERAL INFORMATION:
; APPLICANT: Harvie, Pierrot
; APPLICANT: Paul, Ralph
; APPLICANT: Cudmore, Sally
; APPLICANT: O'Mahony, Daniel J.
; TITLE OF INVENTION: LIPID-COMPRISING DRUG DELIVERY COMPLEXES
; TITLE OF INVENTION: AND METHODS FOR THEIR PRODUCTION
; FILE REFERENCE: 226272005300
; CURRENT APPLICATION NUMBER: US/10/850,873
; CURRENT FILING DATE: 2004-05-20
; PRIOR APPLICATION NUMBER: US/10/136,187
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/287,786
; PRIOR FILING DATE: 2001-04-30
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 28
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Artificial Sequence

```

```
; FEATURE:
; OTHER INFORMATION: Synthetic Construct
; NAME/KEY: VARIANT
; LOCATION: 1
; OTHER INFORMATION: Lysine is attached to dansyl
US-10-850-873-28

Query Match      69.7%; Score 53; DB 17; Length 16;
Best Local Similarity 85.7%; Pred. No. 0.16;
Matches 12; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      4 RKAANAALLPVLLA 17
DB      1 KKKAANAALLPVLLA 14

RESULT 26
US-10-955-656-48
; Sequence 48, Application US/10955656
; Publication No. US20050101762A1
; GENERAL INFORMATION:
; APPLICANT: O'Mahony, Daniel J.
; APPLICANT: Lambkin, Inelda J.
; APPLICANT: Pinilla, Clemencia
; APPLICANT: Houghten, Richard
; TITLE OF INVENTION: CONJUGATES OF MEMBRANE TRANSLOCATING AGENTS AND PHARMACEUTICALLY
; FILE REFERENCE: P26, 481-A USA
; CURRENT APPLICATION NUMBER: US/10/955,656
; CURRENT FILING DATE: 2004-09-30
; PRIOR APPLICATION NUMBER: 10/126,845
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 09/671,089
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/156,246
; PRIOR FILING DATE: 1999-09-27
; NUMBER OF SEQ ID NOS: 119
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 48
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: dansylated membrane translocating peptide
; NAME/KEY: MOD_RES
; LOCATION: (1)..(1)
; OTHER INFORMATION: dansylated
US-10-955-656-48

Query Match      69.7%; Score 53; DB 17; Length 16;
Best Local Similarity 85.7%; Pred. No. 0.16;
Matches 12; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      4 RKAANAALLPVLLA 17
DB      1 KKKAANAALLPVLLA 14

RESULT 27
US-10-136-187-5
; Sequence 5, Application US/10136187
; Publication No. US20030203865A1
; GENERAL INFORMATION:
; APPLICANT: Harvie, Pierrot
; APPLICANT: Paul, Ralph
; APPLICANT: Cudmore, Sally
; APPLICANT: O'Mahony, Daniel J.
; TITLE OF INVENTION: LIPID-COMPRISING DRUG DELIVERY COMPLEXES
; FILE REFERENCE: 226272005300
; CURRENT APPLICATION NUMBER: US/10/136,187
```

```
; CURRENT FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: US 60/287,786
; PRIOR FILING DATE: 2001-04-30
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Construct
US-10-136-187-5

Query Match      69.7%; Score 53; DB 15; Length 20;
Best Local Similarity 85.7%; Pred. No. 0.2;
Matches 12; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      4 RKAANAALLPVLLA 17
DB      1 KKKAANAALLPVLLA 14

RESULT 28
US-10-136-187-9
; Sequence 9, Application US/10136187
; Publication No. US20030203865A1
; GENERAL INFORMATION:
; APPLICANT: Harvie, Pierrot
; APPLICANT: Paul, Ralph
; APPLICANT: Cudmore, Sally
; APPLICANT: O'Mahony, Daniel J.
; TITLE OF INVENTION: LIPID-COMPRISING DRUG DELIVERY COMPLEXES
; FILE REFERENCE: 226272005300
; CURRENT APPLICATION NUMBER: US/10/136,187
; CURRENT FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: US 60/287,786
; PRIOR FILING DATE: 2001-04-30
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Construct
; NAME/KEY: VARIANT
; LOCATION: 1
; OTHER INFORMATION: Lysine is attached to galactose
US-10-136-187-9

Query Match      69.7%; Score 53; DB 15; Length 20;
Best Local Similarity 85.7%; Pred. No. 0.2;
Matches 12; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      4 RKAANAALLPVLLA 17
DB      1 KKKAANAALLPVLLA 14

RESULT 29
US-10-850-873-5
; Sequence 5, Application US/10850873
; Publication No. US20050025821A1
; GENERAL INFORMATION:
; APPLICANT: Harvie, Pierrot
; APPLICANT: Paul, Ralph
; APPLICANT: Cudmore, Sally
; APPLICANT: O'Mahony, Daniel J.
; TITLE OF INVENTION: LIPID-COMPRISING DRUG DELIVERY COMPLEXES
; FILE REFERENCE: 226272005300
```

; CURRENT APPLICATION NUMBER: US/10/850,873  
; CURRENT FILING DATE: 2004-05-20  
; PRIOR APPLICATION NUMBER: US/10/136,187  
; PRIOR FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: US 60/287,786  
; PRIOR FILING DATE: 2001-04-30  
; NUMBER OF SEQ ID NOS: 45  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5  
; LENGTH: 20  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Synthetic Construct  
US-10-850-873-5

Query Match 69.7%; Score 53; DB 17; Length 20;  
Best Local Similarity 85.7%; Pred. No. 0.2;  
Matches 12; Conservative 1; Mismatches 1; Indels 1; Gaps 0;

QY 4 RKAAAVLLPVLLA 17  
DB 1 KKKAAVLLPVLLA 14

RESULT 30  
US-10-850-873-9  
; Sequence 9, Application US/10850873  
; Publication No. US20050025821A1  
; GENERAL INFORMATION:  
; APPLICANT: Harvie, Pierrot  
; APPLICANT: Paul, Ralph  
; APPLICANT: Cudmore, Sally  
; APPLICANT: O'Mahony, Daniel J.  
; TITLE OF INVENTION: LIPID-COMPRISING DRUG DELIVERY COMPLEXES  
; TITLE OF INVENTION: AND METHODS FOR THEIR PRODUCTION  
; FILE REFERENCE: 226272005300  
; CURRENT APPLICATION NUMBER: US/10/850,873  
; CURRENT FILING DATE: 2004-05-20  
; PRIOR APPLICATION NUMBER: US/10/136,187  
; PRIOR FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: US 60/287,786  
; PRIOR FILING DATE: 2001-04-30  
; NUMBER OF SEQ ID NOS: 45  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 9  
; LENGTH: 20  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Synthetic Construct  
; FEATURE:  
; NAME/KEY: VARIANT  
; LOCATION: 1  
; OTHER INFORMATION: Lysine is attached to galactose  
US-10-850-873-9

Query Match 69.7%; Score 53; DB 17; Length 20;  
Best Local Similarity 85.7%; Pred. No. 0.2;  
Matches 12; Conservative 1; Mismatches 1; Indels 1; Gaps 0;

QY 4 RKAAAVLLPVLLA 17  
DB 1 KKKAAVLLPVLLA 14

RESULT 31  
US-10-126-845-109  
; Sequence 109, Application US/10126845  
; Publication No. US20030181367A1  
; GENERAL INFORMATION:  
; APPLICANT: O'Mahony, Daniel J.  
; APPLICANT: Lambkin, Imelda J.

; APPLICANT: Pinilla, Clemencia  
; APPLICANT: Houghten, Richard  
; TITLE OF INVENTION: MEMBRANE TRANSLOCATING PEPTIDE DRUG DELIVERY SYSTEM  
; FILE REFERENCE: E1067/20058  
; CURRENT APPLICATION NUMBER: US/10/126,845  
; CURRENT FILING DATE: 2002-10-15  
; NUMBER OF SEQ ID NOS: 119  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 109  
; LENGTH: 21  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Comprises opioid peptide  
; FEATURE:  
; NAME/KEY: MISC\_FEATURE  
; LOCATION: (1)..(1)  
; OTHER INFORMATION: Linked to biotin-LC  
; FEATURE:  
; NAME/KEY: MISC\_FEATURE  
; LOCATION: (18)..(21)  
; OTHER INFORMATION: D-form amino acid  
; FEATURE:  
; NAME/KEY: MOD\_RES  
; LOCATION: (21)..(21)  
; OTHER INFORMATION: Amidation  
US-10-126-845-109

Query Match 69.7%; Score 53; DB 14; Length 21;  
Best Local Similarity 85.7%; Pred. No. 0.21;  
Matches 12; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 RKAAAVLLPVLLA 17  
DB 1 KKKAAVLLPVLLA 14

RESULT 32  
US-10-126-845-111  
; Sequence 111, Application US/10126845  
; Publication No. US20030181367A1  
; GENERAL INFORMATION:  
; APPLICANT: O'Mahony, Daniel J.  
; APPLICANT: Lambkin, Imelda J.  
; APPLICANT: Pinilla, Clemencia  
; APPLICANT: Houghten, Richard  
; TITLE OF INVENTION: MEMBRANE TRANSLOCATING PEPTIDE DRUG DELIVERY SYSTEM  
; FILE REFERENCE: E1067/20058  
; CURRENT APPLICATION NUMBER: US/10/126,845  
; CURRENT FILING DATE: 2002-10-15  
; NUMBER OF SEQ ID NOS: 119  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 111  
; LENGTH: 21  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Comprises opioid peptide  
; FEATURE:  
; NAME/KEY: MISC\_FEATURE  
; LOCATION: (1)..(1)  
; OTHER INFORMATION: Linked to biotin-LC  
; FEATURE:  
; NAME/KEY: MISC\_FEATURE  
; LOCATION: (2)..(21)  
; OTHER INFORMATION: D-form amino acid  
; FEATURE:  
; NAME/KEY: MOD\_RES  
; LOCATION: (21)..(21)  
; OTHER INFORMATION: Amidation  
US-10-126-845-111

Query Match 69.7%; Score 53; DB 14; Length 21;



```

Best Local Similarity 85.7%; Pred. No. 0.21;
Matches 12; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 4 KKA...VLLPVLLA 17
Db 1 KKA...VLLPVLLA 14

RESULT 33
US-10-955-656-109
; Sequence 109, Application US/10955656
; Publication No. US20050101762A1
; GENERAL INFORMATION:
; APPLICANT: O'Mahony, Daniel J.
; APPLICANT: Lambkin, Imelda J.
; APPLICANT: Pinilla, Clemencia
; APPLICANT: Houghten, Richard
; TITLE OF INVENTION: CONJUGATES OF MEMBRANE TRANSLOCATING AGENTS AND PHARMACEUTICALLY
; FILE REFERENCE: P26,481-A USA
; CURRENT APPLICATION NUMBER: US/10/955,656
; CURRENT FILING DATE: 2004-09-30
; PRIOR APPLICATION NUMBER: 10/126,845
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 09/671,089
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/156,246
; PRIOR FILING DATE: 1999-09-27
; NUMBER OF SEQ ID NOS: 119
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 109
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Comprises opioid peptide
; NAME/KEY: MISC_FEATURE
; LOCATION: (1)..(21)
; OTHER INFORMATION: Linked to biotin-LC
; NAME/KEY: MISC_FEATURE
; LOCATION: (18)..(21)
; OTHER INFORMATION: D-form amino acid
; NAME/KEY: MOD_RES
; LOCATION: (21)..(21)
; OTHER INFORMATION: Amidation
; US-10-955-656-109

Query Match 69.7%; Score 53; DB 17; Length 21;
Best Local Similarity 85.7%; Pred. No. 0.21;
Matches 12; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 4 KKA...VLLPVLLA 17
Db 1 KKA...VLLPVLLA 14

RESULT 34
US-10-955-656-111
; Sequence 111, Application US/10955656
; Publication No. US20050101762A1
; GENERAL INFORMATION:
; APPLICANT: O'Mahony, Daniel J.
; APPLICANT: Lambkin, Imelda J.
; APPLICANT: Pinilla, Clemencia
; APPLICANT: Houghten, Richard
; TITLE OF INVENTION: CONJUGATES OF MEMBRANE TRANSLOCATING AGENTS AND PHARMACEUTICALLY
; FILE REFERENCE: P26,481-A USA
; CURRENT APPLICATION NUMBER: US/10/955,656
; CURRENT FILING DATE: 2004-09-30

```

```

; PRIOR APPLICATION NUMBER: 10/126,845
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 09/671,089
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/156,246
; PRIOR FILING DATE: 1999-09-27
; NUMBER OF SEQ ID NOS: 119
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 111
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Comprises opioid peptide
; NAME/KEY: MISC_FEATURE
; LOCATION: (1)..(1)
; OTHER INFORMATION: Linked to biotin-LC
; NAME/KEY: MISC_FEATURE
; LOCATION: (2)..(21)
; OTHER INFORMATION: D-form amino acid
; NAME/KEY: MOD_RES
; LOCATION: (21)..(21)
; OTHER INFORMATION: Amidation
; US-10-955-656-111

Query Match 69.7%; Score 53; DB 17; Length 21;
Best Local Similarity 85.7%; Pred. No. 0.21;
Matches 12; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 4 KKA...VLLPVLLA 17
Db 1 KKA...VLLPVLLA 14

RESULT 35
US-10-126-845-2
; Sequence 2, Application US/10126845
; Publication No. US20030181367A1
; GENERAL INFORMATION:
; APPLICANT: O'Mahony, Daniel J.
; APPLICANT: Lambkin, Imelda J.
; APPLICANT: Pinilla, Clemencia
; APPLICANT: Houghten, Richard
; TITLE OF INVENTION: MEMBRANE TRANSLOCATING PEPTIDE DRUG DELIVERY SYSTEM
; FILE REFERENCE: E1067/20058
; CURRENT APPLICATION NUMBER: US/10/126,845
; CURRENT FILING DATE: 2002-10-15
; NUMBER OF SEQ ID NOS: 119
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 2
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: membrane translocating peptide
; US-10-126-845-2

Query Match 67.1%; Score 51; DB 14; Length 15;
Best Local Similarity 92.3%; Pred. No. 0.3;
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 5 KKA...VLLPVLLA 17
Db 1 KKA...VLLPVLLA 13

RESULT 36
US-10-126-845-3
; Sequence 3, Application US/10126845
; Publication No. US20030181367A1

```

```
; GENERAL INFORMATION:
; APPLICANT: O'Mahony, Daniel J.
; APPLICANT: Lambkin, Imelda J.
; APPLICANT: Pinilla, Clemencia
; APPLICANT: Houghten, Richard
; TITLE OF INVENTION: MEMBRANE TRANSLOCATING PEPTIDE DRUG DELIVERY SYSTEM
; FILE REFERENCE: E1067/20058
; CURRENT APPLICATION NUMBER: US/10/126,845
; CURRENT FILING DATE: 2002-10-15
; NUMBER OF SEQ ID NOS: 119
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: membrane translocating peptide
; NAME/KEY: MOD RES
; LOCATION: (15)..(15)
; OTHER INFORMATION: linked to FITC-LC
;
US-10-126-845-3
Query Match 67.1%; Score 51; DB 14; Length 15;
Best Local Similarity 92.3%; Pred. No. 0.3;
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 KAAAVLLPVLLA 17
| | | | | | | | | |
Db 1 KAAAVLLPVLLA 13

RESULT 37
US-10-126-845-60
; Sequence 60, Application US/10126845
; Publication No. US20030181367A1
; GENERAL INFORMATION:
; APPLICANT: O'Mahony, Daniel J.
; APPLICANT: Lambkin, Imelda J.
; APPLICANT: Pinilla, Clemencia
; APPLICANT: Houghten, Richard
; TITLE OF INVENTION: MEMBRANE TRANSLOCATING PEPTIDE DRUG DELIVERY SYSTEM
; FILE REFERENCE: E1067/20058
; CURRENT APPLICATION NUMBER: US/10/126,845
; CURRENT FILING DATE: 2002-10-15
; NUMBER OF SEQ ID NOS: 119
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 60
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: D form peptide
; NAME/KEY: MISC FEATURE
; LOCATION: (1)..(15)
; OTHER INFORMATION: D form amino acid
;
US-10-126-845-60
Query Match 67.1%; Score 51; DB 14; Length 15;
Best Local Similarity 92.3%; Pred. No. 0.3;
Matches 12; Conservative 0; Mismatches 0; Gaps 0;

QY 5 KAAAVLLPVLLA 17
| | | | | | | | | |
Db 1 KAAAVLLPVLLA 13

RESULT 38
US-10-126-845-61
; Sequence 61, Application US/10126845
; Publication No. US20030181367A1
; GENERAL INFORMATION:
; APPLICANT: O'Mahony, Daniel J.
; APPLICANT: Lambkin, Imelda J.
; APPLICANT: Pinilla, Clemencia
; APPLICANT: Houghten, Richard
; TITLE OF INVENTION: MEMBRANE TRANSLOCATING PEPTIDE DRUG DELIVERY SYSTEM
; FILE REFERENCE: E1067/20058
; CURRENT APPLICATION NUMBER: US/10/126,845
; CURRENT FILING DATE: 2002-10-15
; NUMBER OF SEQ ID NOS: 119
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 61
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: D form peptide
; NAME/KEY: MISC FEATURE
; LOCATION: (1)..(15)
; OTHER INFORMATION: D form amino acid
;
US-10-126-845-61
Query Match 67.1%; Score 51; DB 14; Length 15;
Best Local Similarity 92.3%; Pred. No. 0.3;
Matches 12; Conservative 0; Mismatches 0; Gaps 0;

QY 5 KAAAVLLPVLLA 17
| | | | | | | | | |
Db 1 KAAAVLLPVLLA 13

RESULT 39
US-10-136-187-2
; Sequence 2, Application US/10136187
; Publication No. US20030203865A1
; GENERAL INFORMATION:
; APPLICANT: Harvie, Pierrot
; APPLICANT: Paul, Ralph
; APPLICANT: Cudmore, Sally
; APPLICANT: O'Mahony, Daniel J.
; TITLE OF INVENTION: LIPID-COMPRISING DRUG DELIVERY COMPLEXES
; TITLE OF INVENTION: AND METHODS FOR THEIR PRODUCTION
; FILE REFERENCE: 22672005300
; CURRENT APPLICATION NUMBER: US/10/136,187
; CURRENT FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: US 60/287,786
; PRIOR FILING DATE: 2001-04-30
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Construct
;
US-10-136-187-2
Query Match 67.1%; Score 51; DB 15; Length 15;
Best Local Similarity 92.3%; Pred. No. 0.3;
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 KAAAVLLPVLLA 17
| | | | | | | | | |
Db 1 KAAAVLLPVLLA 13

RESULT 40
US-10-136-187-6
; Sequence 6, Application US/10136187
; Publication No. US20030203865A1
; GENERAL INFORMATION:
; APPLICANT: Harvie, Pierrot
; APPLICANT: Paul, Ralph
```

```
; APPLICANT: O'Mahony, Daniel J.
; APPLICANT: Lambkin, Imelda J.
; APPLICANT: Pinilla, Clemencia
; APPLICANT: Houghten, Richard
; TITLE OF INVENTION: MEMBRANE TRANSLOCATING PEPTIDE DRUG DELIVERY SYSTEM
; FILE REFERENCE: E1067/20058
; CURRENT APPLICATION NUMBER: US/10/126,845
; CURRENT FILING DATE: 2002-10-15
; NUMBER OF SEQ ID NOS: 119
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 61
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: D form peptide
; NAME/KEY: MISC FEATURE
; LOCATION: (1)..(15)
; OTHER INFORMATION: D form amino acid
;
US-10-126-845-61
Query Match 67.1%; Score 51; DB 14; Length 15;
Best Local Similarity 92.3%; Pred. No. 0.3;
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 KAAAVLLPVLLA 17
| | | | | | | | | |
Db 1 KAAAVLLPVLLA 13

RESULT 39
US-10-136-187-2
; Sequence 2, Application US/10136187
; Publication No. US20030203865A1
; GENERAL INFORMATION:
; APPLICANT: Harvie, Pierrot
; APPLICANT: Paul, Ralph
; APPLICANT: Cudmore, Sally
; APPLICANT: O'Mahony, Daniel J.
; TITLE OF INVENTION: LIPID-COMPRISING DRUG DELIVERY COMPLEXES
; TITLE OF INVENTION: AND METHODS FOR THEIR PRODUCTION
; FILE REFERENCE: 22672005300
; CURRENT APPLICATION NUMBER: US/10/136,187
; CURRENT FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: US 60/287,786
; PRIOR FILING DATE: 2001-04-30
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Construct
;
US-10-136-187-2
Query Match 67.1%; Score 51; DB 15; Length 15;
Best Local Similarity 92.3%; Pred. No. 0.3;
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 KAAAVLLPVLLA 17
| | | | | | | | | |
Db 1 KAAAVLLPVLLA 13

RESULT 40
US-10-136-187-6
; Sequence 6, Application US/10136187
; Publication No. US20030203865A1
; GENERAL INFORMATION:
; APPLICANT: Harvie, Pierrot
; APPLICANT: Paul, Ralph
```

; APPLICANT: Cudmore, Sally  
; APPLICANT: O'Mahony, Daniel J.  
; TITLE OF INVENTION: LIPID-COMPRISING DRUG DELIVERY COMPLEXES  
; TITLE OF INVENTION: AND METHODS FOR THEIR PRODUCTION  
; FILE REFERENCE: 226272005300  
; CURRENT APPLICATION NUMBER: US/10/136,187  
; CURRENT FILING DATE: 2002-09-13  
; PRIOR APPLICATION NUMBER: US 60/287,786  
; PRIOR FILING DATE: 2001-04-30  
; NUMBER OF SEQ ID NOS: 45  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 6  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; FEATURE:  
; NAME/KEY: VARIANT  
; LOCATION: 1  
; OTHER INFORMATION: Lysine is attached to galactose  
US-10-136-187-6

Query Match 67.1%; Score 51; DB 15; Length 15;  
Best Local Similarity 92.3%; Pred. No. 0.3;  
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 KAAAVLLPVLLA 17  
| | | | | | | | | |  
Db 1 KAAAVLLPVLLA 13

## RESULT 41

US-10-136-187-12  
; Sequence 12, Application US/10136187  
; Publication No. US20030203865A1  
; GENERAL INFORMATION:  
; APPLICANT: Harvie, Pierrot  
; APPLICANT: Cudmore, Sally  
; APPLICANT: O'Mahony, Daniel J.  
; TITLE OF INVENTION: LIPID-COMPRISING DRUG DELIVERY COMPLEXES  
; TITLE OF INVENTION: AND METHODS FOR THEIR PRODUCTION  
; FILE REFERENCE: 226272005300  
; CURRENT APPLICATION NUMBER: US/10/136,187  
; CURRENT FILING DATE: 2002-09-13  
; PRIOR APPLICATION NUMBER: US 60/287,786  
; PRIOR FILING DATE: 2001-04-30  
; NUMBER OF SEQ ID NOS: 45  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 12  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Synthetic Construct  
; FEATURE:  
; NAME/KEY: VARIANT  
; LOCATION: 1  
; OTHER INFORMATION: Lysine is attached to DOPE-succinyl  
US-10-136-187-12

Query Match 67.1%; Score 51; DB 15; Length 15;  
Best Local Similarity 92.3%; Pred. No. 0.3;  
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 KAAAVLLPVLLA 17  
| | | | | | | | | |  
Db 1 KAAAVLLPVLLA 13

## RESULT 42

US-10-136-187-13

; Sequence 13, Application US/10136187  
; Publication No. US20030203865A1  
; GENERAL INFORMATION:  
; APPLICANT: Harvie, Pierrot  
; APPLICANT: Paul, Ralph  
; APPLICANT: Cudmore, Sally  
; APPLICANT: O'Mahony, Daniel J.  
; TITLE OF INVENTION: LIPID-COMPRISING DRUG DELIVERY COMPLEXES  
; TITLE OF INVENTION: AND METHODS FOR THEIR PRODUCTION  
; FILE REFERENCE: 226272005300  
; CURRENT APPLICATION NUMBER: US/10/136,187  
; CURRENT FILING DATE: 2002-09-13  
; PRIOR APPLICATION NUMBER: US 60/287,786  
; PRIOR FILING DATE: 2001-04-30  
; NUMBER OF SEQ ID NOS: 45  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 13  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Synthetic Construct  
US-10-136-187-13

Query Match 67.1%; Score 51; DB 15; Length 15;  
Best Local Similarity 92.3%; Pred. No. 0.3;  
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 KAAAVLLPVLLA 17  
| | | | | | | | | |  
Db 1 KAAAVLLPVLLA 13

## RESULT 43

US-10-136-187-31  
; Sequence 31, Application US/10136187  
; Publication No. US20030203865A1  
; GENERAL INFORMATION:  
; APPLICANT: Harvie, Pierrot  
; APPLICANT: Paul, Ralph  
; APPLICANT: Cudmore, Sally  
; APPLICANT: O'Mahony, Daniel J.  
; TITLE OF INVENTION: LIPID-COMPRISING DRUG DELIVERY COMPLEXES  
; TITLE OF INVENTION: AND METHODS FOR THEIR PRODUCTION  
; FILE REFERENCE: 226272005300  
; CURRENT APPLICATION NUMBER: US/10/136,187  
; CURRENT FILING DATE: 2002-09-13  
; PRIOR APPLICATION NUMBER: US 60/287,786  
; PRIOR FILING DATE: 2001-04-30  
; NUMBER OF SEQ ID NOS: 45  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 31  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Synthetic Construct  
; FEATURE:  
; NAME/KEY: VARIANT  
; LOCATION: 1  
; OTHER INFORMATION: Lysine is attached to des-Proline  
; NAME/KEY: VARIANT  
; LOCATION: 15  
; OTHER INFORMATION: Serine is attached to galactose  
US-10-136-187-31

Query Match 67.1%; Score 51; DB 15; Length 15;  
Best Local Similarity 92.3%; Pred. No. 0.3;  
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 KAAAVLLPVLLA 17  
| | | | | | | | | |

Db 1 KKA0AVLLPVLLA 13

## RESULT 44

US-10-136-187-33  
; Sequence 33, Application US/10136187  
; Publication No. US20030203865A1  
; GENERAL INFORMATION:  
; APPLICANT: Harvie, Pierrot  
; APPLICANT: Paul, Ralph  
; APPLICANT: Cudmore, Sally  
; APPLICANT: O'Mahony, Daniel J.  
; TITLE OF INVENTION: LIPID-COMPRISING DRUG DELIVERY COMPLEXES  
; FILE REFERENCE: 226272005300  
; CURRENT APPLICATION NUMBER: US/10/136,187  
; CURRENT FILING DATE: 2002-09-13  
; PRIOR APPLICATION NUMBER: US 60/287,786  
; PRIOR FILING DATE: 2001-04-30  
; NUMBER OF SEQ ID NOS: 45  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 33  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Synthetic Construct  
; NAME/KEY: VARIANT  
; LOCATION: 1  
; OTHER INFORMATION: Lysine is attached to Cholesteryl-succinyl  
US-10-136-187-33

Query Match 67.1%; Score 51; DB 15; Length 15;  
Best Local Similarity 92.3%; Pred. No. 0.3;  
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 KKA0AVLLPVLLA 17

Db 1 KKA0AVLLPVLLA 13

## RESULT 45

US-10-136-187-34  
; Sequence 34, Application US/10136187  
; Publication No. US20030203865A1  
; GENERAL INFORMATION:  
; APPLICANT: Harvie, Pierrot  
; APPLICANT: Paul, Ralph  
; APPLICANT: Cudmore, Sally  
; APPLICANT: O'Mahony, Daniel J.  
; TITLE OF INVENTION: LIPID-COMPRISING DRUG DELIVERY COMPLEXES  
; FILE REFERENCE: 226272005300  
; CURRENT APPLICATION NUMBER: US/10/136,187  
; CURRENT FILING DATE: 2002-09-13  
; PRIOR APPLICATION NUMBER: US 60/287,786  
; PRIOR FILING DATE: 2001-04-30  
; NUMBER OF SEQ ID NOS: 45  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 34  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Synthetic Construct  
; NAME/KEY: VARIANT  
; LOCATION: 1  
; OTHER INFORMATION: Lysine is attached to DOPE-succinyl  
US-10-136-187-34

Query Match 67.1%; Score 51; DB 15; Length 15;

Best Local Similarity 92.3%; Pred. No. 0.3;  
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 KKA0AVLLPVLLA 17

Db 1 KKA0AVLLPVLLA 13

## RESULT 46

US-10-136-187-35  
; Sequence 35, Application US/10136187  
; Publication No. US20030203865A1  
; GENERAL INFORMATION:  
; APPLICANT: Harvie, Pierrot  
; APPLICANT: Paul, Ralph  
; APPLICANT: Cudmore, Sally  
; APPLICANT: O'Mahony, Daniel J.  
; TITLE OF INVENTION: LIPID-COMPRISING DRUG DELIVERY COMPLEXES  
; FILE REFERENCE: 226272005300  
; CURRENT APPLICATION NUMBER: US/10/136,187  
; CURRENT FILING DATE: 2002-09-13  
; PRIOR APPLICATION NUMBER: US 60/287,786  
; PRIOR FILING DATE: 2001-04-30  
; NUMBER OF SEQ ID NOS: 45  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 35  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Synthetic Construct  
; NAME/KEY: VARIANT  
; LOCATION: 1  
; OTHER INFORMATION: Lysine is attached to Cholesteryl-succinyl  
US-10-136-187-35

Query Match 67.1%; Score 51; DB 15; Length 15;  
Best Local Similarity 92.3%; Pred. No. 0.3;  
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 KKA0AVLLPVLLA 17

Db 1 KKA0AVLLPVLLA 13

## RESULT 47

US-10-136-187-36  
; Sequence 36, Application US/10136187  
; Publication No. US20030203865A1  
; GENERAL INFORMATION:  
; APPLICANT: Harvie, Pierrot  
; APPLICANT: Paul, Ralph  
; APPLICANT: Cudmore, Sally  
; APPLICANT: O'Mahony, Daniel J.  
; TITLE OF INVENTION: LIPID-COMPRISING DRUG DELIVERY COMPLEXES  
; FILE REFERENCE: 226272005300  
; CURRENT APPLICATION NUMBER: US/10/136,187  
; CURRENT FILING DATE: 2002-09-13  
; PRIOR APPLICATION NUMBER: US 60/287,786  
; PRIOR FILING DATE: 2001-04-30  
; NUMBER OF SEQ ID NOS: 45  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 36  
; LENGTH: 15  
; TYPE: PRT  
; ORGANISM: Artificial Sequence

```
; FEATURE:
; OTHER INFORMATION: Synthetic Construct
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 1
; OTHER INFORMATION: Lysine is attached to DSPE-PEG5K-succinyl
US-10-136-187-36

Query Match      67.1%; Score 51; DB 15; Length 15;
Best Local Similarity 92.3%; Pred. No. 0.3;
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 KAAAVLLPVLLA 17
DB 1 KAAAVLLPVLLA 13

RESULT 48
US-10-136-187-37
; Sequence 37, Application US/10136187
; Publication No. US20030203865A1
; GENERAL INFORMATION:
; APPLICANT: Harvie, Pierrot
; APPLICANT: Paul, Ralph
; APPLICANT: Cudmore, Sally
; APPLICANT: O'Mahony, Daniel J.
; TITLE OF INVENTION: LIPID-COMPRISING DRUG DELIVERY COMPLEXES
; TITLE OF INVENTION: AND METHODS FOR THEIR PRODUCTION
; FILE REFERENCE: 226272005300
; CURRENT APPLICATION NUMBER: US/10/136,187
; CURRENT FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: US 60/287,786
; PRIOR FILING DATE: 2001-04-30
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 37
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Construct
; NAME/KEY: VARIANT
; LOCATION: 1
; OTHER INFORMATION: Lysine is attached to DMPE-PEG5K-succinyl
US-10-136-187-37

Query Match      67.1%; Score 51; DB 15; Length 15;
Best Local Similarity 92.3%; Pred. No. 0.3;
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 KAAAVLLPVLLA 17
DB 1 KAAAVLLPVLLA 13

RESULT 49
US-10-136-187-38
; Sequence 38, Application US/10136187
; Publication No. US20030203865A1
; GENERAL INFORMATION:
; APPLICANT: Harvie, Pierrot
; APPLICANT: Paul, Ralph
; APPLICANT: Cudmore, Sally
; APPLICANT: O'Mahony, Daniel J.
; TITLE OF INVENTION: LIPID-COMPRISING DRUG DELIVERY COMPLEXES
; TITLE OF INVENTION: AND METHODS FOR THEIR PRODUCTION
; FILE REFERENCE: 226272005300
; CURRENT APPLICATION NUMBER: US/10/136,187
; CURRENT FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: US 60/287,786
; PRIOR FILING DATE: 2001-04-30
; NUMBER OF SEQ ID NOS: 45
```

```
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 38
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Construct
; NAME/KEY: VARIANT
; LOCATION: 1
; OTHER INFORMATION: Lysine is attached to DSPE-PEG5K-succinyl
US-10-136-187-38

Query Match      67.1%; Score 51; DB 15; Length 15;
Best Local Similarity 92.3%; Pred. No. 0.3;
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 KAAAVLLPVLLA 17
DB 1 KAAAVLLPVLLA 13

RESULT 50
US-10-136-187-39
; Sequence 39, Application US/10136187
; Publication No. US20030203865A1
; GENERAL INFORMATION:
; APPLICANT: Harvie, Pierrot
; APPLICANT: Paul, Ralph
; APPLICANT: Cudmore, Sally
; APPLICANT: O'Mahony, Daniel J.
; TITLE OF INVENTION: LIPID-COMPRISING DRUG DELIVERY COMPLEXES
; TITLE OF INVENTION: AND METHODS FOR THEIR PRODUCTION
; FILE REFERENCE: 226272005300
; CURRENT APPLICATION NUMBER: US/10/136,187
; CURRENT FILING DATE: 2002-09-13
; PRIOR APPLICATION NUMBER: US 60/287,786
; PRIOR FILING DATE: 2001-04-30
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 39
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Construct
; NAME/KEY: VARIANT
; LOCATION: 1
; OTHER INFORMATION: Lysine is attached to DMPE-PEG5K-succinyl
US-10-136-187-39

Query Match      67.1%; Score 51; DB 15; Length 15;
Best Local Similarity 92.3%; Pred. No. 0.3;
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 KAAAVLLPVLLA 17
DB 1 KAAAVLLPVLLA 13

Search completed: June 2, 2005, 01:43:52
Job time: 103.344 secs
```

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 2, 2005, 01:26:16 ; Search time 13.2812 Seconds  
(without alignments)  
123.157 Million cell updates/sec

Title: US-10-764-235-24

Perfect score: 76

Sequence: 1 KKKRKAANVLLPVLLA 17

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 750 summaries

Database :

PIR 79:\*\*

1: piri:\*\*

2: pir2:\*\*

3: pir3:\*\*

4: pir4:\*\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query %	Match	Length	ID	Description
1	42	55.3	297	2	AF3275	transporter, dme f
2	42	55.3	527	2	G69635	PTS arbutin-like e
3	41	53.9	528	2	P64580	hypothetical prote
4	41	53.9	582	2	I48673	matrix metalloprot
5	41	53.9	582	2	I38028	matrix metalloprot
6	41	53.9	582	2	I84471	matrix metalloprot
7	40	52.6	310	2	T26710	hypothetical prote
8	40	52.6	389	2	H82825	transport protein
9	40	52.6	534	2	F90031	PTS system, arbuti
10	40	52.6	863	2	T38016	importin beta-1 su
11	40	52.6	980	2	H84632	probable receptor-
12	39	51.3	162	2	G84363	hypothetical prote
13	39	51.3	304	2	JCS847	chitinase (EC 3.2.
14	39	51.3	312	2	S78242	cytochrome c-type
15	39	51.3	318	2	D87506	glycosyl transfera
16	39	51.3	437	2	T42653	hypothetical prote
17	39	51.3	560	2	A70628	probable acid-CoA
18	39	51.3	1089	2	S53978	PSE1 protein - yea
19	39	51.3	1467	2	T48162	hypothetical prote
20	38	50.0	57	2	D82395	hypothetical prote
21	38	50.0	71	2	C82807	30S ribosomal prot
22	38	50.0	71	2	G83572	30S ribosomal prot
23	38	50.0	169	2	A12391	hypothetical prote
24	38	50.0	287	2	T05461	hypothetical prote
25	38	50.0	354	2	G75619	probable hemin ABC
26	38	50.0	411	2	H75376	hypothetical prote
27	38	50.0	548	2	G87227	acyl-CoA synthase
28	38	50.0	627	2	T04562	hypothetical prote
29	38	50.0	722	2	T46560	swi2 protein - fls

hypothetical prote  
hypothetical prote  
hypothetical prote  
glycerol dehydroge  
glycerol dehydroge  
hypothetical prote  
hypothetical prote  
signal peptidase I  
CDPdiacylglycerol-  
glycerol-3-phospha  
hypothetical prote  
conserved hypothet  
outer membrane pro  
probable membrane  
hypothetical prote  
hypothetical prote  
adenylate cyclase  
probable retroelem  
probable retroelem  
genome polyprotein  
BPUF1 protein - hu  
hypothetical prote  
hypothetical prote  
hypothetical prote  
hypothetical prote  
probable secreted  
hypothetical prote  
hypothetical prote  
hypothetical prote  
hypothetical prote  
I3 protein - vacu  
L3L protein - vari  
K3L protein - vari  
hypothetical prote  
glutamate racemase  
glutamate racemase  
hypothetical prote  
hypothetical prote  
probable membrane  
hypothetical prote  
2-dehydro-3-deoxy-  
hypothetical prote  
hypothetical prote  
exopolysaccharide  
hypothetical prote  
exoQ-like protein  
argininosuccinate  
carboxypeptidase (  
hypothetical prote  
hypothetical prote  
myeloperoxidase (E  
regulatory protein  
competence regulat  
two-component sens  
xeroderma pigmento  
chromosome segrega  
xanthine dehydroge  
copia-like polypro  
hypothetical prote  
protein kinase - f  
reinoblastoma bin  
hypothetical prote  
outer membrane pro  
outer membrane por

103	35	46.1	91	2	G69750	hypothetical prote	176	35	46.1	926	2	B37271	A-alpha Y 3 protei
104	35	46.1	131	2	AH1257	diacylglycerol kin	177	35	46.1	961	2	E86245	hypothetical prote
105	35	46.1	131	2	AD1820	diacylglycerol kin	178	35	46.1	1230	2	S47466	cellulose 1,4-beta
106	35	46.1	143	2	G84742	hypothetical prote	179	35	46.1	1356	2	T16754	hypothetical prote
107	35	46.1	155	2	G69987	hypothetical prote	180	34.5	45.4	662	2	E83201	conserved hypotet
108	35	46.1	168	2	D29016	cell division inhi	181	34	44.7	42	2	A37907	thyrotropin beta c
109	35	46.1	168	2	C90209	hypothetical prote	182	34	44.7	95	2	AD1168	hypothetical prote
110	35	46.1	172	2	A69818	hypothetical prote	183	34	44.7	104	2	E86263	Fl3K23.22 protein
111	35	46.1	206	2	JC7320	K562 cell-derived	184	34	44.7	119	2	C82918	ribosomal protein
112	35	46.1	234	2	T34686	probable integral	185	34	44.7	141	2	S34629	hypothetical prote
113	35	46.1	248	1	S50179	indole-3-glycerol-	186	34	44.7	142	2	S58247	probable membrane
114	35	46.1	248	2	C40635	indole-3-glycerol-	187	34	44.7	155	1	TI2M1	trypsin/factor XII
115	35	46.1	254	2	AC3018	conserved hypotet	188	34	44.7	179	2	H87667	hypothetical prote
116	35	46.1	254	2	E98266	hypothetical prote	189	34	44.7	186	2	G81319	phosphoheptose iso
117	35	46.1	257	2	A32283	MHC class II histo	190	34	44.7	189	2	E69823	conserved hypotet
118	35	46.1	289	2	T37340	30K DNA binding ph	191	34	44.7	191	2	E86664	hypothetical prote
119	35	46.1	271	2	F83188	phosphatidate cyti	192	34	44.7	203	2	AC3483	thiol:disulfide in
120	35	46.1	271	2	JC4832	phosphatidate cyti	193	34	44.7	209	2	A71313	probable V-type AT
121	35	46.1	288	2	T05954	transcription fact	194	34	44.7	213	2	T02871	germin-like protei
122	35	46.1	288	2	S33714	ribosomal protein	195	34	44.7	225	2	G86748	hypothetical prote
123	35	46.1	288	2	I51803	TAXRBB107 - human	196	34	44.7	227	2	T47186	hypothetical prote
124	35	46.1	308	2	JC7125	epidermal growth f	197	34	44.7	235	2	E72805	gp48 protein - Myc
125	35	46.1	316	2	B97622	thioredoxin reduct	198	34	44.7	252	2	H81209	hypothetical prote
126	35	46.1	317	2	F70673	probable mutri pro	199	34	44.7	252	2	H72469	hypothetical prote
127	35	46.1	319	2	JC5402	vitamin D receptor	200	34	44.7	275	2	H82929	ATP synthase A cha
128	35	46.1	324	2	AH2844	thioredoxin reduct	201	34	44.7	277	2	C70017	4-hydroxybenzoyl-C
129	35	46.1	325	2	AF2865	conserved hypotet	202	34	44.7	282	2	E81448	pantoate-beta-alan
130	35	46.1	330	2	C91233	2-keto-3-deoxy-D-g	203	34	44.7	282	2	E70553	hypothetical prote
131	35	46.1	330	2	C86080	2-keto-3-deoxy-D-g	204	34	44.7	293	2	T37227	probable sugar tra
132	35	46.1	330	2	S40853	probable 2-keto-3-	205	34	44.7	297	2	E96002	probable sugar upt
133	35	46.1	333	2	E97447	hypothetical prote	206	34	44.7	303	2	B83336	hypothetical prote
134	35	46.1	335	2	C97336	transcription regu	207	34	44.7	303	2	T36337	probable membrane
135	35	46.1	348	2	G87604	thioredoxin reduct	208	34	44.7	310	2	D72615	hypothetical prote
136	35	46.1	350	2	E70788	hypothetical prote	209	34	44.7	319	2	F41337	endo-1,4-beta-xyla
137	35	46.1	352	2	S70972	bfpE protein - Esc	210	34	44.7	323	2	A58892	NADH2 dehydrogenas
138	35	46.1	356	2	S34649	genome polypeptin	211	34	44.7	324	2	C79719	enoyl-acyl carrier
139	35	46.1	374	2	S75459	hypothetical prote	212	34	44.7	346	2	S77025	nitrilase (EC 3.5.
140	35	46.1	390	2	S46441	aspartate carbamoy	213	34	44.7	346	2	JC5715	G protein-coupled
141	35	46.1	398	2	JQ0113	2-keto-3-deoxygluc	214	34	44.7	349	2	AH1764	B. subtilis probab
142	35	46.1	402	1	S70969	bundle-forming pil	215	34	44.7	352	2	AF1389	B. subtilis probab
143	35	46.1	407	2	T36979	probable integral	216	34	44.7	355	2	F95979	probable transcrip
144	35	46.1	411	2	T35562	probable integral	217	34	44.7	359	2	JC7280	cytokine receptor-
145	35	46.1	423	2	AI2896	FAD-dependent mono	218	34	44.7	368	2	D87271	general secretion
146	35	46.1	423	2	C97672	probable transmem	219	34	44.7	371	2	T22894	hypothetical prote
147	35	46.1	426	2	T35550	probable amino aci	220	34	44.7	374	2	AG0937	hypothetical prote
148	35	46.1	460	2	AI0689	probable integral	221	34	44.7	387	2	F84392	hypothetical prote
149	35	46.1	469	2	T35526	elastin binding pr	222	34	44.7	389	2	T72554	hypothetical prote
150	35	46.1	486	2	A89927	stromelysin 3 (EC	223	34	44.7	390	2	AF0842	multidrug resistan
151	35	46.1	488	2	S13423	hypothetical prote	224	34	44.7	405	2	A46699	translation initia
152	35	46.1	513	2	J39534	transcription inic	225	34	44.7	410	2	T51212	related to integra
153	35	46.1	528	2	JN0445	probable exported	226	34	44.7	412	2	D72364	conserved hypotet
154	35	46.1	533	2	AG1018	probable exported	227	34	44.7	420	2	T05877	hypothetical prote
155	35	46.1	533	2	C64706	phosphate permease	228	34	44.7	421	2	D97078	MDR-type permease,
156	35	46.1	533	2	A71812	probable phosphate	229	34	44.7	429	2	G70528	probable arsaA prot
157	35	46.1	536	2	AH0913	probable exported	230	34	44.7	434	2	G98235	hypothetical prote
158	35	46.1	538	2	S22409	D-alanyl-D-alanine	231	34	44.7	445	2	AE3050	two component sens
159	35	46.1	589	2	T50698	probable mandelonl	232	34	44.7	452	2	A83734	PTS system, n-acet
160	35	46.1	589	2	T19216	hypothetical prote	233	34	44.7	457	2	AI0955	probable PFS syste
161	35	46.1	592	2	H82031	probable biotin-la	234	34	44.7	457	2	C81430	probable periplasm
162	35	46.1	598	2	D84242	hypothetical prote	235	34	44.7	458	2	S77605	probable membrane
163	35	46.1	611	2	T19217	hypothetical prote	236	34	44.7	463	2	H72365	heat shock protein
164	35	46.1	613	2	T47483	receptor like prot	237	34	44.7	465	2	AE0931	two-component syst
165	35	46.1	644	2	JC5119	anti-mullerian hor	238	34	44.7	473	2	H84550	probable obtusifol
166	35	46.1	677	2	S65573	phosphate-specific	239	34	44.7	482	2	JC7583	basic helix-loop-h
167	35	46.1	741	2	AD3271	hypothetical prote	240	34	44.7	491	1	T36811	probable integral
168	35	46.1	749	2	AD0345	probable phosphate	241	34	44.7	493	1	IJBQCP	P-cadherin - bovin
169	35	46.1	806	1	E70131	endopeptidase la (	242	34	44.7	498	2	A75509	ABC transporter, A
170	35	46.1	809	2	T41645	probable spindle p	243	34	44.7	504	2	T16526	hypothetical prote
171	35	46.1	810	2	C72785	probable aldehyde	244	34	44.7	513	2	T37180	probable membrane
172	35	46.1	821	1	IJMSCP	P-cadherin precurs	245	34	44.7	515	2	H84013	hypothetical prote
173	35	46.1	841	2	T14520	probable S-recepto	246	34	44.7	516	2	S28339	dhlc protein - Xan
174	35	46.1	844	1	T05179	S-receptor kinase	247	34	44.7	519	2	S39893	rncC protein - Rho
175	35	46.1	878	2	AD0664	probable exported	248	34	44.7	531	2	B96966	beta-mannanase Man



249	34	44.7	543	2	F81665	ADP, ATP carrier p	322	33	43.4	222	2	E71507	probable L4 riboso
250	34	44.7	546	2	AF2268	type II site-speci	323	33	43.4	222	2	G75078	hypothetical prote
251	34	44.7	548	2	T51035	hypothetical prote	324	33	43.4	224	2	C86571	L4 ribosomal prote
252	34	44.7	550	2	C86704	conserved hypotet	325	33	43.4	224	2	H72055	ribosomal protein
253	34	44.7	556	2	T46842	K+-transporting AT	326	33	43.4	232	2	T22698	hypothetical prote
254	34	44.7	557	2	H97351	K+-transporting AT	327	33	43.4	234	1	S15102	eosinophil major b
255	34	44.7	604	2	H95974	msba-like sacchari	328	33	43.4	243	2	A13422	superoxide diamuta
256	34	44.7	604	2	S60182	ATP-binding transp	329	33	43.4	246	2	F70062	hypothetical prote
257	34	44.7	608	2	F83397	probable peptidase	330	33	43.4	250	2	AD2704	nitrogen fixation
258	34	44.7	639	2	B86587	DNA helicase [impo	331	33	43.4	250	2	D97486	nitrogen fixation
259	34	44.7	639	2	G72038	ATP-dependent heli	332	33	43.4	251	2	H97207	hypothetical secre
260	34	44.7	639	2	T32269	hypothetical prote	333	33	43.4	255	2	T36012	probable secreted
261	34	44.7	644	2	E88487	protein C13B9.4 [i	334	33	43.4	257	2	A82369	phosphorybosilform
262	34	44.7	654	2	D43689	hypothetical prote	335	33	43.4	264	2	T30391	hypothetical prote
263	34	44.7	669	2	T43866	phosphate transpor	336	33	43.4	273	2	T51010	hypothetical prote
264	34	44.7	716	2	D69585	conserved hypotet	337	33	43.4	275	2	AF1160	motility protein (
265	34	44.7	744	2	F75625	serine proteinase,	338	33	43.4	275	2	AF1519	hypothetical protei
266	34	44.7	766	2	A85198	hypothetical prote	339	33	43.4	281	2	D70682	hypothetical prote
267	34	44.7	780	2	T27669	hypothetical prote	340	33	43.4	282	2	G84672	hypothetical prote
268	34	44.7	790	2	T30081	hypothetical prote	341	33	43.4	284	2	H84096	transcription regu
269	34	44.7	806	2	T25614	hypothetical prote	342	33	43.4	287	2	C82593	ABC transporter ph
270	34	44.7	815	2	T15749	hypothetical prote	343	33	43.4	287	2	T51011	hypothetical prote
271	34	44.7	819	2	T45690	receptor-like prot	344	33	43.4	290	2	AE0677	probable regulator
272	34	44.7	829	1	IJHUCP	cadherin 3 precurs	345	33	43.4	292	2	S52888	SUP1 protein - yea
273	34	44.7	901	2	S65161	hypothetical prote	346	33	43.4	295	2	H87612	TPR domain protein
274	34	44.7	977	2	H86428	hypothetical prote	347	33	43.4	297	2	F83491	probable transcrip
275	34	44.7	977	2	I52657	seizure-related pr	348	33	43.4	298	2	T37251	homeobox protein c
276	34	44.7	1035	2	A86177	hypothetical prote	349	33	43.4	299	2	T08704	hypothetical prote
277	34	44.7	1056	1	DJAD51	DNA-directed DNA p	350	33	43.4	300	2	T44730	CDP-diacylglycerol
278	34	44.7	1056	1	WMAD12	DNA-directed DNA p	351	33	43.4	304	2	A83928	sugar transport sy
279	34	44.7	1061	1	DJAD12	DNA-directed DNA p	352	33	43.4	304	2	AF2615	conserved hypotet
280	34	44.7	1115	2	A47541	protein kinase INE	353	33	43.4	308	2	F87283	methionyl-tRNA for
281	34	44.7	1456	2	T15961	hypothetical prote	354	33	43.4	308	2	D70875	probable PE protei
282	34	44.7	1698	2	S51869	probable membrane	355	33	43.4	313	2	E97397	hypothetical prote
283	34	44.7	1952	2	T48814	hypothetical prote	356	33	43.4	314	2	A87621	hypothetical prote
284	34	44.7	2009	2	S45764	SEC7 protein - yea	357	33	43.4	316	2	T36909	probable dihydroepi
285	34	44.7	3071	2	T45584	hypothetical prote	358	33	43.4	324	2	E95048	enoyl-(acyl-carrie
286	33.5	44.1	278	2	A13633	hypothetical prote	359	33	43.4	329	2	G75395	ABC transporter, p
287	33.5	44.1	376	2	E70680	probable proB - My	360	33	43.4	329	2	F69849	probable heme A fa
288	33.5	44.1	773	2	G96571	hypothetical prote	361	33	43.4	330	2	S32773	alpha-antigen - My
289	33.5	44.1	878	2	T37978	multiple BRCCT doma	362	33	43.4	335	2	S73392	type I restriction
290	33	43.4	52	2	B60407	monocyte adherence	363	33	43.4	335	2	B87590	hypothetical prote
291	33	43.4	70	2	G81023	30S ribosomal prot	364	33	43.4	336	2	D71474	probable muramoyl-
292	33	43.4	71	2	S45608	light-harvesting p	365	33	43.4	346	2	A95324	hypothetical prote
293	33	43.4	94	1	WSWLC1	E5 protein - pygmy	366	33	43.4	353	2	AH3116	transcription regu
294	33	43.4	112	2	S62929	hypothetical prote	367	33	43.4	353	2	E98170	transcription regu
295	33	43.4	113	2	A64914	hypothetical prote	368	33	43.4	356	2	C47116	trifoliotoxin synth
296	33	43.4	119	2	A85654	probable diacygly	369	33	43.4	358	2	H36891	transfer complex p
297	33	43.4	119	2	D90793	probable diacygly	370	33	43.4	361	2	A53216	prostaglandin E2 r
298	33	43.4	128	2	C83161	hypothetical prote	371	33	43.4	362	2	T52038	probable phytochel
299	33	43.4	133	2	T35848	probable secreted	372	33	43.4	365	2	I38748	prostaglandin E re
300	33	43.4	144	2	B95905	conserved hypotet	373	33	43.4	365	2	S51315	prostaglandin E re
301	33	43.4	145	2	S73871	type I restriction	374	33	43.4	367	2	JC2056	histone H2A.1 - ra
302	33	43.4	157	2	A72558	probable ribosomal	375	33	43.4	368	2	I80811	probable inosine-5
303	33	43.4	158	2	A61069	replication protei	376	33	43.4	370	2	D86957	hypothetical prote
304	33	43.4	176	2	S52914	virion protein j13	377	33	43.4	371	2	T01708	prostaglandin recep
305	33	43.4	179	2	H69524	mutator protein Mu	378	33	43.4	374	2	I38747	prostaglandin recep
306	33	43.4	183	2	S52904	virion protein j13	379	33	43.4	374	2	S51317	prostaglandin E re
307	33	43.4	184	2	T33006	hypothetical prote	380	33	43.4	375	1	S72812	IMP dehydrogenase-
308	33	43.4	185	2	S52903	virion protein j13	381	33	43.4	379	2	A83225	ABC transporter, m
309	33	43.4	189	2	S52902	virion protein j13	382	33	43.4	384	2	JC4802	alkaline proteinas
310	33	43.4	189	2	S52907	virion protein j13	383	33	43.4	386	2	E97196	probable glycosyl
311	33	43.4	191	2	S52905	virion protein j13	384	33	43.4	388	2	S51316	prostaglandin E re
312	33	43.4	199	2	C84311	hypothetical prote	385	33	43.4	388	2	I38750	prostaglandin recep
313	33	43.4	199	2	H75531	cytochrome c-type	386	33	43.4	389	2	AF0920	uroporphyrinogen I
314	33	43.4	201	2	E84187	hypothetical prote	387	33	43.4	390	2	S43375	prostaglandin E re
315	33	43.4	204	2	H96491	hypothetical prote	388	33	43.4	390	2	S51313	prostaglandin E re
316	33	43.4	206	2	JC4268	fibroblast growth	389	33	43.4	390	2	T28083	hypothetical prote
317	33	43.4	212	2	T05721	germin-like protei	390	33	43.4	393	2	S51318	prostaglandin E re
318	33	43.4	218	2	T43438	hypothetical prote	391	33	43.4	393	2	B96780	hypothetical prote
319	33	43.4	220	2	A28456	histone H1.10 - ch	392	33	43.4	394	2	E82572	ABC transporter so
320	33	43.4	222	2	T11942	ribosomal protein	393	33	43.4	395	2	T09373	hypothetical prote
321	33	43.4	222	2	C81661	ribosomal protein	394	33	43.4	396	2	B75290	hypothetical prote

395	33	43.4	396	2	B83461	hypothetical prote	468	33	43.4	658	2	B81225	glutathione-regula
396	33	43.4	397	2	S66518	proteinase-activat	469	33	43.4	658	2	C81997	probable glutathio
397	33	43.4	402	2	T04348	endosperm specific	470	33	43.4	658	2	D81099	membrane-bound lyc
398	33	43.4	407	1	A49845	5-aminolevulinat	471	33	43.4	658	2	D81842	probable membrane
399	33	43.4	407	2	B32306	cytochrome P450 C	472	33	43.4	705	2	T35165	probable integral
400	33	43.4	411	2	B53216	prostaglandin E2 r	473	33	43.4	710	2	F86778	DNA topoisomerase
401	33	43.4	411	2	C30091	capA protein - Bac	474	33	43.4	715	2	B07411	probable moey prot
402	33	43.4	411	2	S09833	hypothetical prote	475	33	43.4	731	2	A83160	conserved hypothet
403	33	43.4	412	2	E97279	TPr-repeat-contain	476	33	43.4	748	2	T35818	probable integral
404	33	43.4	421	1	S70433	zona pellucida gly	477	33	43.4	751	2	T15403	hypothetical prote
405	33	43.4	422	2	C70518	probable nant prot	478	33	43.4	753	2	G70184	methyl-accepting c
406	33	43.4	425	2	S51319	prostaglandin E re	479	33	43.4	765	2	A36332	regulatory protein
407	33	43.4	425	2	C95001	MesJ/Ycf62 family	480	33	43.4	777	2	T04768	hypothetical prote
408	33	43.4	425	2	B97873	conserved hypothet	481	33	43.4	796	2	A90541	hypothetical prote
409	33	43.4	428	2	D75582	hypothetical prote	482	33	43.4	805	2	A56199	transcription fact
410	33	43.4	435	2	H70033	maltodextrin trans	483	33	43.4	815	2	T15402	hypothetical prote
411	33	43.4	440	2	T24837	hypothetical prote	484	33	43.4	827	1	A36895	endopeptidase La (
412	33	43.4	443	2	B86849	polysaccharide bios	485	33	43.4	944	2	T18911	hypothetical prote
413	33	43.4	444	1	AJBSQU	glutamate-ammonia	486	33	43.4	955	2	S56649	pyruvate, phosphat
414	33	43.4	445	2	C72695	probable amidase A	487	33	43.4	956	1	S53297	pyruvate, phosphat
415	33	43.4	448	2	B87635	transcription regu	488	33	43.4	993	1	PIVXTA	RNA i protein - to
416	33	43.4	456	2	E83750	gluconate permease	489	33	43.4	1048	2	T19148	hypothetical prote
417	33	43.4	457	2	D70901	probable fmu prote	490	33	43.4	1130	2	T19148	hypothetical prote
418	33	43.4	461	2	JE0030	nitrogenase (EC 1.	491	33	43.4	1234	2	T30160	hypothetical prote
419	33	43.4	461	2	B81080	drug resistance tr	492	33	43.4	1247	1	MMHUND	nidogen precursor
420	33	43.4	461	2	B81862	probable integral	493	33	43.4	1270	2	T30339	deRNA adenosine de
421	33	43.4	467	2	G87564	aminopeptidase, pr	494	33	43.4	1430	2	T34516	hypothetical prote
422	33	43.4	468	2	B95895	probable protein (	495	33	43.4	1487	1	EDBEE1	immediate-early pr
423	33	43.4	468	2	A12956	hypothetical prote	496	33	43.4	1487	1	EDBEF6	155K transcritpion
424	33	43.4	471	2	B38637	Ras inhibitor (clo	497	33	43.4	1508	2	E87696	glutamate synthase
425	33	43.4	472	2	S77364	DNA damage-inducib	498	33	43.4	1549	2	D86338	protein F5W15.18 (
426	33	43.4	487	1	B64795	ybdS protein - Esc	499	33	43.4	1721	1	I38902	retinoblastoma bin
427	33	43.4	487	2	S75338	NADH2 dehydrogenas	500	33	43.4	2870	2	A35548	319K protein ndvB
428	33	43.4	487	2	C80560	probable membrane	501	33	43.4	4092	1	S38128	helic2 protein - mo
429	33	43.4	487	2	C90710	hypothetical prote	502	33	43.4	4836	2	T14346	herc2 protein - mo
430	33	43.4	487	2	E98326	hypothetical prote	503	33	43.4	4859	2	S74173	ryanodine receptor
431	33	43.4	488	2	B95332	hypothetical prote	504	33	43.4	4872	2	S27272	ryanodine receptor
432	33	43.4	492	2	S46225	ammonium transport	505	32.5	42.8	82	2	A12710	30S ribosomal prot
433	33	43.4	500	2	I56502	4-aminobutyrate tr	506	32.5	42.8	109	2	AG3358	hypothetical prote
434	33	43.4	501	2	A80578	citrate carrier (i	507	32.5	42.8	167	1	H71235	hypothetical prote
435	33	43.4	503	2	AC3296	proline/betaine tr	508	32.5	42.8	236	2	F83025	hypothetical prote
436	33	43.4	504	2	S52048	deoxyribodipyrimid	509	32.5	42.8	257	2	AF2592	2-deoxyribose-5-ph
437	33	43.4	513	2	AR2432	sodium/solute symp	510	32.5	42.8	259	2	H97374	2-deoxyribose-5-ph
438	33	43.4	517	1	QRS8UB	sfuB protein - Ser	511	32.5	42.8	434	2	A83197	citrate synthase (
439	33	43.4	520	2	E89044	protein B0238.9 [i	512	32.5	42.8	491	2	C46230	RNA-binding protei
440	33	43.4	532	2	B82102	membrane-bound lyc	513	32.5	42.8	536	2	B46230	RNA-binding protei
441	33	43.4	534	2	F72679	hypothetical prote	514	32.5	42.8	541	2	A48717	glutamate formimid
442	33	43.4	536	1	C64728	probable membrane	515	32.5	42.8	615	2	S24761	Cpo 61.1 protein -
443	33	43.4	536	2	AH0515	probable ABC trans	516	32.5	42.8	616	2	A46230	RNA-binding protei
444	33	43.4	536	2	G85488	probable transport	517	32.5	42.8	655	2	E81157	cell division prot
445	33	43.4	536	2	G90637	probable transport	518	32.5	42.8	655	2	E81948	probable ATP-depen
446	33	43.4	544	1	S41389	penton protein (ii	519	32.5	42.8	784	2	T51759	[glutamate-ammonia
447	33	43.4	544	2	A47710	metalloproteinase	520	32	42.1	79	2	A83851	Cpo 61.1 protein -
448	33	43.4	562	2	I40227	neutrol proteinase	521	32	42.1	81	2	T02667	RNA-binding protei
449	33	43.4	562	2	T49788	related to merzoi	522	32	42.1	87	2	T29724	cell division prot
450	33	43.4	583	2	T07848	pectinesterase (EC	523	32	42.1	95	2	S72597	probable ATP-depen
451	33	43.4	589	2	T40252	hypothetical prote	524	32	42.1	97	1	GECH	[glutamate-ammonia
452	33	43.4	590	1	A45621	leishmanolysin (EC	525	32	42.1	111	2	B40076	cytochrome-c oxida
453	33	43.4	592	2	I49239	vesicle transport	526	32	42.1	115	2	B49094	methylmalonyl-CoA
454	33	43.4	598	2	C64215	leishmanolysin (EC	527	32	42.1	117	2	D72650	hypothetical prote
455	33	43.4	599	2	B42049	GTP-binding membra	528	32	42.1	118	2	H81443	diacylglycerol kin
456	33	43.4	599	2	A44951	leishmanolysin (EC	529	32	42.1	120	2	F71126	hypothetical prote
457	33	43.4	600	2	C83142	hypothetical prote	530	32	42.1	123	2	AD0039	diacylglycerol kin
458	33	43.4	602	1	PL0221	leishmanolysin (EC	531	32	42.1	124	2	I50095	MHC class II beta
459	33	43.4	604	1	HMNZCD	hemagglutinin - ca	532	32	42.1	129	2	B83053	secretion protein
460	33	43.4	607	2	S49017	hemagglutinin - ca	533	32	42.1	129	2	F82116	flagellar protein
461	33	43.4	615	2	T34392	hypothetical prote	534	32	42.1	137	2	F95239	conserved domain p
462	33	43.4	615	2	AH2448	proteinase [import	535	32	42.1	137	2	H98103	hypothetical prote
463	33	43.4	619	2	G75321	ABC transporter, A	536	32	42.1	138	2	A29479	thyrotropin beta c
464	33	43.4	625	2	T21051	hypothetical prote	537	32	42.1	140	2	I46880	T-cell receptor be
465	33	43.4	638	2	A36929	virulence regulato	538	32	42.1	142	2	E69945	hypothetical prote
466	33	43.4	656	2	JN0666	dnak-type molecula	539	32	42.1	148	2	G83615	3-dehydroquinat d
467	33	43.4	657	2	T48228	probable protein k	540	32	42.1	154	2	E87132	very hypothetical

541	32	42.1	162	2	F81413	hypothetical prote	614	32	42.1	298	2	H71501	hypothetical prote
542	32	42.1	163	2	H75388	hypothetical prote	615	32	42.1	301	2	B82235	conserved hypotet
543	32	42.1	165	2	S24968	ecp2 protein - fun	616	32	42.1	304	2	AF0730	probable membrane
544	32	42.1	169	2	A83134	conserved hypotet	617	32	42.1	306	2	AF1142	ABC transporter (b
545	32	42.1	171	1	HSUR1E	histone H1, gonada	618	32	42.1	306	2	AF1500	ABC transporter (b
546	32	42.1	172	2	F83696	hypothetical prote	619	32	42.1	307	2	TI4814	hypothetical prote
547	32	42.1	173	2	D83355	gluconokinase PA23	620	32	42.1	309	2	T33259	hypothetical prote
548	32	42.1	173	2	S44607	COPF5.5 protein -	621	32	42.1	309	2	T22402	hypothetical prote
549	32	42.1	175	2	B44102	di-N-acetylchitobi	622	32	42.1	312	2	C82298	tRNA pseudouridine
550	32	42.1	176	2	AF0684	bacteriophage host	623	32	42.1	312	2	C71243	hypothetical prote
551	32	42.1	176	2	A97500	hypothetical prote	624	32	42.1	313	2	C75208	sugar transport sy
552	32	42.1	176	2	AE2718	conserved hypotet	625	32	42.1	315	2	TI7340	hypothetical prote
553	32	42.1	178	2	H81285	probable flagellar	626	32	42.1	315	2	RWRBB	T-cell receptor be
554	32	42.1	179	2	H72697	hypothetical prote	627	32	42.1	320	2	A82449	conserved hypotet
555	32	42.1	180	2	F75502	purine/pyrimidine	628	32	42.1	322	2	S75428	UDP-N-acetylglucos
556	32	42.1	182	2	H87383	hypothetical prote	629	32	42.1	324	2	AB3316	thioredoxin-disulf
557	32	42.1	189	2	F97496	hypothetical prote	630	32	42.1	325	2	TI7307	hypothetical prote
558	32	42.1	189	2	AF3127	hypothetical prote	631	32	42.1	326	2	E82764	hypothetical prote
559	32	42.1	192	2	A13437	colicin v producti	632	32	42.1	328	2	G75551	hypothetical prote
560	32	42.1	194	2	I50710	fibroblast growth	633	32	42.1	329	2	B65078	hypothetical prote
561	32	42.1	200	2	H81126	conserved hypotet	634	32	42.1	329	2	B85950	probable oxidoredu
562	32	42.1	200	2	C81894	probable integral	635	32	42.1	333	1	D69812	ferrichrome ABC tr
563	32	42.1	201	2	T44629	hypothetical prote	636	32	42.1	343	2	B72507	hypothetical prote
564	32	42.1	206	2	H87187	conserved hypotet	637	32	42.1	345	2	T34998	probable transmemb
565	32	42.1	207	2	F81040	cytochrome C4 NMB1	638	32	42.1	346	2	AC1331	proteinases homolo
566	32	42.1	208	2	C70899	probable guanylate	639	32	42.1	346	2	AC1702	weakly proteinases
567	32	42.1	213	2	T30618	hypothetical prote	640	32	42.1	348	2	T07669	cyclin al-type, m1
568	32	42.1	214	2	AD3279	hypothetical prote	641	32	42.1	348	2	T36890	probable iron-side
569	32	42.1	214	2	T44107	conserved hypotet	642	32	42.1	348	2	E84588	hypothetical prote
570	32	42.1	214	2	G89763	conserved hypotet	643	32	42.1	349	2	A88197	protein ZK1127.4 l
571	32	42.1	218	2	G87574	rhomboid family pr	644	32	42.1	349	2	T36152	probable transpos
572	32	42.1	220	2	S35075	class II histocomp	645	32	42.1	349	2	A97007	probable membrane
573	32	42.1	228	2	E86849	glutamate ABC tran	646	32	42.1	357	2	E90404	hypothetical prote
574	32	42.1	228	2	AF0728	Prop effector [imp	647	32	42.1	363	2	A38430	trans-activation-r
575	32	42.1	229	2	C82117	conserved hypotet	648	32	42.1	364	2	H69831	acetyl-CoA C-acety
576	32	42.1	229	2	T06204	hypothetical prote	649	32	42.1	364	2	B84055	phospho-2-dehydro
577	32	42.1	230	2	I50097	MHC class II beta	650	32	42.1	366	1	G01420	TAR RNA-binding pr
578	32	42.1	230	2	JQ1880	hypothetical 25.2K	651	32	42.1	370	2	I51199	activin beta B sub
579	32	42.1	231	2	I50467	MHC class II beta	652	32	42.1	371	2	B64044	rod shape-determin
580	32	42.1	232	2	E90946	Prop effector [imp	653	32	42.1	372	2	I39868	sac operon regulat
581	32	42.1	233	2	S47625	aspartate carbamoy	654	32	42.1	372	2	AG2049	permease protein o
582	32	42.1	236	2	B83144	conserved hypotet	655	32	42.1	384	1	I38890	dual specificity p
583	32	42.1	237	2	AG30602	T-cell receptor be	656	32	42.1	384	2	AB1212	fatty acid efflux
584	32	42.1	238	2	AG3512	attachment mediati	657	32	42.1	386	2	T06484	aspartate carbamoy
585	32	42.1	239	2	S31033	gene 88 protein -	658	32	42.1	388	2	C95175	membrane protein l
586	32	42.1	240	2	D72809	Sp88 protein - Myc	659	32	42.1	388	2	G75375	hypothetical prote
587	32	42.1	243	2	AE1260	hypothetical prote	660	32	42.1	388	2	D98041	conserved hypotet
588	32	42.1	246	2	I50094	MHC class II beta	661	32	42.1	389	2	H70841	probable amia prot
589	32	42.1	247	2	I50093	MHC class II beta	662	32	42.1	389	2	E49905	protein secretion
590	32	42.1	247	2	I50466	MHC class II beta	663	32	42.1	392	1	A54227	phosphoribosylglyc
591	32	42.1	247	2	I50465	MHC class II beta	664	32	42.1	392	2	AG0741	phosphoribosylglyc
592	32	42.1	248	1	HSUR1P	histone H1, gonada	665	32	42.1	392	2	C85797	phosphoribosylglyc
593	32	42.1	248	2	S35076	class II histocomp	666	32	42.1	392	2	G90948	phosphoribosylglyc
594	32	42.1	252	2	H90476	hypothetical prote	667	32	42.1	393	2	AE0216	phosphoribosylglyc
595	32	42.1	260	1	F64149	nitrate transport	668	32	42.1	393	2	B86189	protein T25N20.9 l
596	32	42.1	265	1	OTRZ3M	cytochrome-c oxida	669	32	42.1	394	2	F81789	multidrug resistan
597	32	42.1	265	1	OTWT3M	cytochrome-c oxida	670	32	42.1	394	2	T36852	probable integral
598	32	42.1	265	2	S20801	cytochrome-c oxida	671	32	42.1	398	2	TI40074	hypothetical zf-c3
599	32	42.1	266	2	AB2270	hypothetical prote	672	32	42.1	399	2	TI9115	hypothetical prote
600	32	42.1	267	2	A46122	homeotic protein H	673	32	42.1	400	2	S32804	beta-3-adrenergic
601	32	42.1	269	2	I46884	T-cell receptor be	674	32	42.1	400	2	A41679	beta-3-adrenergic
602	32	42.1	270	2	T43039	hypothetical prote	675	32	42.1	400	2	A53281	beta-3-adrenergic
603	32	42.1	277	1	D35115	indole-3-glycerol-	676	32	42.1	400	2	S23888	hypothetical prote
604	32	42.1	277	2	S03240	hypothetical prote	677	32	42.1	403	2	C70385	hypothetical prote
605	32	42.1	280	2	E83476	probable transmemb	678	32	42.1	404	2	T36239	probable integral
606	32	42.1	285	2	AH0529	phosphatidate cyti	679	32	42.1	405	2	S65459	beta-3-adrenergic
607	32	42.1	286	2	AH2884	hypothetical prote	680	32	42.1	405	2	T52307	serine/threonine p
608	32	42.1	286	2	E97660	hypothetical prote	681	32	42.1	406	2	T47674	hypothetical prote
609	32	42.1	289	2	S35022	nolT protein - Rhi	682	32	42.1	418	2	G02953	beta-3-adrenergic
610	32	42.1	291	2	B95316	probable ABC trans	683	32	42.1	420	2	C89184	hypothetical prote
611	32	42.1	295	2	E70022	transcription regu	684	32	42.1	425	2	G91104	probable oxidoredu
612	32	42.1	296	2	C69635	transcription repr	685	32	42.1	427	2	G83361	ArSB protein PA227
613	32	42.1	297	2	D69419	phosphate ABC tran	686	32	42.1	427	2	A43453	A-kinase anchor pr

687 32 42.1 427 2 A84826  
688 32 42.1 429 1 B41902  
689 32 42.1 429 1 C41903  
690 32 42.1 430 2 G89962  
691 32 42.1 431 2 AD1237  
692 32 42.1 431 2 AD1599  
693 32 42.1 433 2 T36122  
694 32 42.1 435 2 G69779  
695 32 42.1 435 2 E70711  
696 32 42.1 435 2 T29593  
697 32 42.1 436 2 JC4841  
698 32 42.1 436 2 G84024  
699 32 42.1 437 2 AF2330  
700 32 42.1 439 2 C89890  
701 32 42.1 439 2 JC2532  
702 32 42.1 441 2 C70702  
703 32 42.1 443 2 T27877  
704 32 42.1 444 2 T27234  
705 32 42.1 445 2 A97873  
706 32 42.1 446 2 AE1023  
707 32 42.1 446 2 S56352  
708 32 42.1 446 2 A91267  
709 32 42.1 446 2 F86107  
710 32 42.1 449 2 I52630  
711 32 42.1 449 2 A36303  
712 32 42.1 449 2 A34625  
713 32 42.1 449 2 A10909  
714 32 42.1 455 1 WQEBST  
715 32 42.1 456 2 S62331  
716 32 42.1 458 2 H83881  
717 32 42.1 458 2 D96928  
718 32 42.1 460 2 A31897  
719 32 42.1 460 2 I51837  
720 32 42.1 460 2 S09508  
721 32 42.1 460 2 A24325  
722 32 42.1 460 2 S39715  
723 32 42.1 460 2 I64080  
724 32 42.1 461 2 G83602  
725 32 42.1 461 2 B82510  
726 32 42.1 462 2 T34984  
727 32 42.1 465 2 T26195  
728 32 42.1 467 2 T48743  
729 32 42.1 471 2 H87284  
730 32 42.1 471 2 BMH03  
731 32 42.1 472 1 F70031  
732 32 42.1 473 2 T17260  
733 32 42.1 473 2 AF0132  
734 32 42.1 475 2 B75409  
735 32 42.1 477 2 C72658  
736 32 42.1 478 2 S26353  
737 32 42.1 479 2 A26048  
738 32 42.1 479 2 D87644  
739 32 42.1 485 2 C75460  
740 32 42.1 490 2 D87644  
741 32 42.1 491 2 C83206  
742 32 42.1 491 2 E95356  
743 32 42.1 497 2 AG1768  
744 32 42.1 498 2 E69033  
745 32 42.1 500 2 T19743  
746 32 42.1 502 2 T22045  
747 32 42.1 518 2 G84488  
748 32 42.1 519 2 AH0064  
749 32 42.1 519 2 AE4811  
750 32 42.1 521 2 AE2226

## ALIGNMENTS

RESULT 1  
AF3275  
transposor, dme family BMEI0187 [imported] - Brucella melitensis (strain 16M)  
C:Species: Brucella melitensis

C:Date: 01-Feb-2002 #sequence\_revision 01-Feb-2002 #text\_change 09-Jul-2004  
C:Accession: AF3275  
R:DelVecchio, V.G.; Kapatral, V.; Redkar, R.J.; Patra, G.; Patra, G.; Los, T.; Ivanova, .; Mazur, M.; Goltzman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letes, Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002  
A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis  
A:Reference number: AD3252; PMID:11756688  
A:Accession: AF3275  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-297 <KUR>  
A:Cross-references: UNIPROT:Q8YJ99; GB:AE008917; PIDN:AAL51369.1; PID:gl7982069; GSPDB:  
A:Experimental source: strain 16M  
C:Genetics:  
A:Gene: BMEI0187  
A:Map position: 1  
Query Match 55.3%; Score 42; DB 2; Length 297;  
Best Local Similarity 81.8%; Pred. No. 19;  
Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
QY 7 AAVALLPVLLA 17  
DB 160 AAALLIPVLLA 170  
RESULT 2  
G69635  
PTS arbutin-like enzyme IIBC component glvC - Bacillus subtilis  
C:Species: Bacillus subtilis  
C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 09-Jul-2004  
C:Accession: G69635  
R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte, C.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch, A.; Ehrlich, S.D.; Emmeron, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.; Nature 390, 249-256, 1997  
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galie, iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F. Koester, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois, A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maue, Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetell Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadate, Y.; Sato, T.; Scanlon A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Sero, akeuchi, M.; Tanakoshi, A.; Tanaka, T.; Terpsira, P.; Tognoni, A.; Tosato, V.; Uchiyama T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, T.; A:Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.  
A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.  
A:Reference number: A69580; MUID:98044033; PMID:9384377  
A:Accession: G69635  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-527 <KUN>  
A:Cross-references: UNIPROT:P54715; GB:Z99108; GB:AL009126; NID:g2633055; PIDN:CAB12649  
A:Experimental source: strain 168  
C:Genetics:  
A:Gene: glvC  
C:Superfamily: phosphotransferase system glucose-specific enzyme II, factor II; phospho  
Query Match 55.3%; Score 42; DB 2; Length 527;  
Best Local Similarity 50.0%; Pred. No. 33;  
Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;  
QY 1 KKKRKAALVLPVLL 16  
DB 300 KKKKLVAGLLIPVTL 315  
RESULT 3  
F64580  
hypothetical protein HP0486 - Helicobacter pylori (strain 26695)  
C:Species: Helicobacter pylori  
C:Date: 09-Aug-1997 #sequence\_revision 09-Aug-1997 #text\_change 09-Jul-2004  
C:Accession: F64580

R.Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R.D.; Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khalak, H.G.; Glodek, A.; McKenney, J.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watthey, L. Nature 388, 539-547, 1997

A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser, C.

A:Title: The complete genome sequence of the gastric pathogen *Helicobacter pylori*.

A:Reference number: A64520; MUID:197394467; PMID:9252185

A:Accession: F64580

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-528 <TOM>

A:Cross-references: UNIPROT:Q25230; GB:AE000563; GB:AE000511; NID:G2313593; PIDN:RAD0758

C:Superfamily: *Helicobacter pylori* hypothetical protein HP0209

Query Match 53.9%; Score 41; DB 2; Length 528;  
Best Local Similarity 61.1%; Pred. No. 48;  
Matches 11; Conservative 2; Mismatches 1; Indels 4; Gaps 1;

QY 2 KKKKAAAVL-----LPVL 15  
||||| ||||| |||||  
DB 4 KKKKVAALLKRFLLPL 21

RESULT 4

I48673

matrix metalloproteinase (EC 3.4.24.-) membrane type precursor - mouse

C:Species: *Mus musculus* (house mouse)

C>Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 22-Jun-1999

C:Accession: I48673

R.Okada, A.; Bellocq, J.P.; Rouyer, N.; Chenard, M.P.; Rio, M.C.; Chambon, P.; Basset, P. Proc. Natl. Acad. Sci. U.S.A. 92, 2730-2734, 1995

A:Title: Membrane-type matrix metalloproteinase (MT-MMP) gene is expressed in stromal cells.

A:Reference number: I38046; MUID:95224014; PMID:7708715

A:Accession: I48673

A:Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: mRNA

A:Residues: 1-582 <RES>

A:Cross-references: EMBL:X83536; NID:G804999; PIDN:CAA58520.1; PID:G805000

C:Superfamily: interstitial collagenase; hemopexin repeat homology; matrix metalloproteinase

C:Keywords: hydrolase; metalloproteinase; zinc; zymogen

F:1-23/Domain: signal sequence #status predicted <SIG>

F:24-97/Domain: activation peptide #status predicted <PRO>

F:161-284/Domain: matrix metalloproteinase homology <MMP>

F:313-508/Domain: hemopexin repeat homology <PXN>

F:93-239,243,249/Binding site: zinc, catalytic (Cys, His, His) (inhibited) #status predicted

F:239,243,249/Binding site: zinc, catalytic (His) (active) #status predicted

F:240/Active site: Glu #status predicted

Query Match 53.9%; Score 41; DB 2; Length 582;  
Best Local Similarity 81.8%; Pred. No. 52;  
Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 6 AAAAVLLPVLL 16  
:|||||:  
DB 538 SAAAVLPLVLL 548

RESULT 5

I38028

matrix metalloproteinase 14 (EC 3.4.24.-) membrane type precursor - human

N:Alternate names: matrix metalloproteinase MMP14; membrane type matrix metalloproteinase

C:Species: *Homo sapiens* (man)

C>Date: 17-May-1996 #sequence\_revision 17-May-1996 #text\_change 09-Jul-2004

C:Accession: I38028; G02274; I38046; S78011; S45341; S71384

R.Will, H.; Hinzmann, B.

Eur. J. Biochem. 231, 602-608, 1995

A:Title: cDNA sequence and mRNA tissue distribution of a novel human matrix metalloproteinase

A:Reference number: I38028; MUID:95377289; PMID:7649159

A:Accession: I38028

A:Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: mRNA

A:Residues: 1-582 <WILL>

A:Cross-references: UNIPROT:P50281; EMBL:Z48481; NID:G963053; PIDN:CAA88372.1; PID:G9630

R.Luo, G.

submitted to the EMBL Data Library, November 1995

A:Reference number: H00963

A:Accession: G02274

A:Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: mRNA

A:Residues: 1-582 <LUO>

A:Cross-references: EMBL:U41078; NID:G1127836; PIDN:AAA83770.1; PID:G1127837

R.Okada, A.; Bellocq, J.P.; Rouyer, N.; Chenard, M.P.; Rio, M.C.; Chambon, P.; Basset, P. Proc. Natl. Acad. Sci. U.S.A. 92, 2730-2734, 1995

A:Title: Membrane-type matrix metalloproteinase (MT-MMP) gene is expressed in stromal cells.

A:Reference number: I38046; MUID:95224014; PMID:7708715

A:Accession: I38046

A:Status: translated from GB/EMBL/DBDJ

A:Molecule type: mRNA

A:Residues: 1-7,'S',9-582 <OKA>

A:Cross-references: EMBL:X83535; NID:G804993; PIDN:CAA58519.1; PID:G804994

R.Seiki, M.

submitted to the EMBL Data Library, January 1994

A:Reference number: S78011

A:Accession: S78011

A:Molecule type: mRNA

A:Residues: 1-7,'S',9-337,'K',339-582 <SEI>

A:Cross-references: EMBL:D28512; NID:G793762; PIDN:BAA05519.1; PID:G793763

R.Sato, H.; Takino, T.; Okada, Y.; Cao, J.; Shinagawa, A.; Yamamoto, E.; Seiki, M. Nature 370, 61-65, 1994

A:Title: A matrix metalloproteinase expressed on the surface of invasive tumour cells.

A:Reference number: S45341; MUID:94286011; PMID:8015608

A:Accession: S45341

A:Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-7,'S',9-188,'R',190,'A',192-267,'K',269-272,'HY',275,'P',277-285,'KQ',288,289

R.Sato, H.; Kinoshita, T.; Takino, T.; Nakayama, K.; Seiki, M. FEBS Lett. 393, 101-104, 1996

A:Title: Activation of a recombinant membrane type 1-matrix metalloproteinase (MT1-MMP)

A:Reference number: S71384; MUID:96397540; PMID:8804434

A:Accession: S71384

A:Molecule type: protein

A:Residues: 112-116 <SAW>

C:Genetics: GDB:MMP14; MT1-MMP

A:Gene: GDB:MMP14; MT1-MMP

A:Cross-references: GDB:375731; OMIM:600754

A:Map position: 14q11-14q12

C:Superfamily: interstitial collagenase; hemopexin repeat homology; matrix metalloproteinase

C:Keywords: glycoprotein; hydrolase; metalloproteinase; zinc; zymogen

F:1-23/Domain: signal sequence #status predicted <SIG>

F:24-97/Domain: activation peptide #status predicted <PRO>

F:161-284/Domain: matrix metalloproteinase homology <MMP>

F:98-582/Product: matrix metalloproteinase 14 membrane type #status predicted <MAT>

F:285-313/Domain: hinge #status predicted <HNG>

F:314-508/Domain: hemopexin repeat homology <PXN>

F:539-582/Domain: transmembrane #status predicted <TM>

F:93-239,243,249/Binding site: zinc, catalytic (Cys, His, His) (inhibited) #status predicted

F:130/Binding site: carboxylate (Asn) (covalent) #status predicted

F:239,243,249/Binding site: zinc, catalytic (His) (active) #status predicted

F:240/Active site: Glu #status predicted

F:319-508/Disulfide bonds: #status predicted

Query Match 53.9%; Score 41; DB 2; Length 582;  
Best Local Similarity 81.8%; Pred. No. 52;  
Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 6 AAAAVLLPVLL 16  
:|||||:  
DB 538 SAAAVLPLVLL 548

RESULT 6

I84471

matrix metalloproteinase (EC 3.4.24.-) membrane type precursor - rat

N:Alternate names: membrane-type metalloproteinase

C:Species: *Rattus norvegicus* (Norway rat)

C;Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 09-Jul-2004  
 C;Accession: I84471; I61946  
 R;Okada, A.; Bellocq, J.P.; Rouyer, N.; Chenard, M.P.; Rio, M.C.; Chambon, P.; Basset, P.  
 Proc. Natl. Acad. Sci. U.S.A. 92, 2730-2734, 1995  
 A;Title: Membrane-type matrix metalloproteinase (MT-MMP) gene is expressed in stromal cells  
 A;Reference number: I38046; MUID:95224014; PMID:7708715  
 A;Accession: I84471  
 A;Status: preliminary; translated from GB/EMBL/DBJ  
 A;Molecule type: mRNA  
 A;Residues: 1-582 <RES>  
 A;Cross-references: UNIPROT:Q10739; EMBL:X83537; NID:g805012; PIDN:CAA58521.1; PID:g805012  
 A;Accession: I61946  
 A;Status: preliminary; translated from GB/EMBL/DBJ  
 A;Molecule type: mRNA  
 A;Residues: 1-67, 'M', 69-254, 'A', 256-582 <RES>  
 A;Cross-references: EMBL:X91785; NID:g1001926; PIDN:CAA62897.1; PID:g1001927  
 C;Genetics:  
 A;Gene: mt-mmp  
 C;Superfamily: interstitial collagenase; hemopexin repeat homology; matrix metalloproteinase  
 C;Keywords: hydrolase; metalloproteinase; zinc; zymogen  
 F;1-23/Domain: signal sequence #status predicted <SIG>  
 F;24-97/Domain: activation peptide #status predicted <PRO>  
 F;61-284/Domain: matrix metalloproteinase homology <MMP>  
 F;313-508/Domain: hemopexin repeat homology <PXN>  
 F;93-239, 243-249/Binding site: zinc, catalytic (Cys, His, His, His) (inhibited) #status  
 F;239,243,249/Binding site: zinc, catalytic (His) (active) #status predicted  
 F;240/Active site: Glu #status predicted

Query Match 53.9%; Score 41; DB 2; Length 582;  
 Best Local Similarity 81.8%; Pred. No. 52;  
 Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 6 AAAAVALLPVLL 16  
 :|:|:|:|:|:|  
 Db 538 SAAAVLPVLL 548

RESULT 7  
 T26710  
 hypothetical protein Y38H6C.2 - Caenorhabditis elegans  
 C;Species: Caenorhabditis elegans  
 C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
 C;Accession: T26710  
 R;White, S.  
 submitted to the EMBL Data Library, September 1998  
 A;Reference number: Z20255  
 A;Accession: T26710  
 A;Status: preliminary; translated from GB/EMBL/DBJ  
 A;Molecule type: DNA  
 A;Residues: 1-310 <WIL>  
 A;Cross-references: UNIPROT:Q9XX39; EMBL:AL031630; PIDN:CAA21000.1; GSPDB:GNO00023; CESP:  
 C;Genetics:  
 A;Gene: CESP:Y38H6C.2  
 A;Map position: 5  
 A;Introns: 38/2; 130/2; 171/3; 209/3; 242/2

Query Match 52.6%; Score 40; DB 2; Length 310;  
 Best Local Similarity 61.5%; Pred. No. 43;  
 Matches 8; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 5 KAAAVALLPVLL 17  
 ||:|:|:|:|  
 Db 171 KACLAILPILTA 183

RESULT 8  
 H82825  
 transport protein XF0281 [imported] - Xylella fastidiosa (strain 9a5c)  
 C;Species: Xylella fastidiosa  
 C;Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 09-Jul-2004  
 C;Accession: H82825  
 R;anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen

Nature 406, 151-157, 2000  
 A;Title: The genome sequence of the plant pathogen Xylella fastidiosa.  
 A;Reference number: A8515; MUID:20365717; PMID:10910347  
 A;Note: For a complete list of authors see reference number A59328 below  
 A;Accession: H82825  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-389 <SIM>  
 A;Cross-references: UNIPROT:Q9PGL7; GB:AE003881; GB:AE003849; NID:g9105093; PIDN:AAF830  
 A;Experimental source: strain 9a5c  
 R;Simpson, A.J.G.; Reinach, P.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Canarigo, L.E.A.; Carraro, D.M.; Carrier, I.; as-Neto, E.; Docena, C.; El-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S.  
 submitted to GenBank, June 2000  
 A;Authors: Ferreira, V.C.A.; Perro, J.A.; Praga, J.S.; Franca, S.C.; Franco, M.C.; Froh  
 J.D.; Junqueira, M.B.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laig  
 chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, A.  
 A;Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.  
 , F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.  
 Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasa  
 A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silve  
 M.; Teuhako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.;  
 A;Reference number: A59328  
 A;Contents: annotation  
 C;Genetics:  
 A;Gene: XF0281

Query Match 52.6%; Score 40; DB 2; Length 389;  
 Best Local Similarity 75.0%; Pred. No. 53;  
 Matches 9; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 6 AAAVALLPVLL 17  
 |:|:|:|:|  
 Db 66 ATQAVLPVLL 77

RESULT 9  
 F90031  
 PTS system, arbutin-like IBC component [imported] - Staphylococcus aureus (strain N315)  
 C;Species: Staphylococcus aureus  
 C;Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 09-Jul-2004  
 C;Accession: F90031  
 R;Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Ogu  
 ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaico, C.; Sekimizu, K.;  
 C.; Shiba, T.; Hattori, M.; Ogawara, N.; Hayashi, H.; Hiramatsu, K.  
 Lancet 357, 1225-1240, 2001  
 A;Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.  
 A;Reference number: A89758; MUID:21311952; PMID:11418146  
 A;Accession: F90031  
 A;Status: preliminary  
 A;Molecule type: DNA  
 A;Residues: 1-534 <KUR>  
 A;Cross-references: UNIPROT:Q99RV0; GB:BA000018; PID:gl3702123; PIDN:BAB43415.1; GSPDB:  
 C;Experimental source: strain N315  
 C;Genetics:  
 A;Gene: glvC  
 C;Superfamily: phosphotransferase system glucose-specific enzyme II, factor II; phospho

Query Match 52.6%; Score 40; DB 2; Length 534;  
 Best Local Similarity 50.0%; Pred. No. 71;  
 Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 KKKRKAATAVALLPVLL 16  
 |:|:|:|:|  
 Db 298 KENRKKVAALLPATL 313

RESULT 10  
 T38016  
 importin beta-1 subunit - fission yeast (Schizosaccharomyces pombe)  
 C;Species: Schizosaccharomyces pombe  
 C;Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004  
 C;Accession: T38016

R.Oliver, K.; Harris, D.; Barrell, B.G.; Rajandream, M.A.; Wood, V.  
submitted to the EMBL Data Library, August 1995  
A:Reference number: Z21762  
A:Accession: T38016  
A:Status: preliminary; translated from GB/EMBL/DBSJ  
A:Molecule type: DNA  
A:Residues: 1-863 <OLI>  
A:CROSS-references: UNIPROT:O13864; EMBL:Z98532; PIDN:CAB11082.1; GSPDB:GN00066; SPDB:SF  
A:Experimental source: strain 972h-; cosmid c1B1  
C:Genetics:  
A:Gene: SPDB:SPAC1B1.03c  
A:Map position: 1

Query Match 52.6%; Score 40; DB 2; Length 863;  
Best Local Similarity 75.0%; Pred. No. 1.1e+02;  
Matches 9; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 5 KAAAVLLPVLL 16  
DB 316 RAAADILPVLL 327

RESULT 11  
H84632  
probable receptor-like protein kinase [imported] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 09-Jul-2004  
A:Accession: H84632  
R:Lin, X.; Kaul, S.; Rounslev, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;  
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;  
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.  
Nature 402, 761-768, 1999  
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
A:Reference number: A84420; MUID:20083487; PMID:10617197  
A:Accession: H84632  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-980 <STO>  
A:CROSS-references: UNIPROT:Q9ZUI0; GB:AE002093; NID:g4115373; PIDN:AAD03374.1; GSPDB:GN  
C:Genetics:  
A:Gene: At2g24130  
A:Map position: 2

Query Match 52.6%; Score 40; DB 2; Length 980;  
Best Local Similarity 56.2%; Pred. No. 1.2e+02;  
Matches 9; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 2 KKKKAAAVLLPVLL 17  
DB 586 KKKKYPVLLPVLLS 601

RESULT 12  
G84363  
hypothetical protein ggt [imported] - Halobacterium sp. NRC-1  
C:Species: Halobacterium sp. NRC-1  
C:Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 09-Jul-2004  
A:Accession: G84363  
R:Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.  
; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jabl  
Jung, K.H.; Alam, M.; Freitas, T.  
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000  
A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omar, A.D.; Ebhardt, H.; Lowe, T.M.; Li  
A:Title: Genome sequence of Halobacterium species NRC-1.  
A:Reference number: A84160; MUID:20504483; PMID:11016950  
A:Accession: G84363  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-162 <STO>  
A:CROSS-references: UNIPROT:Q9HNE5; GB:AE004437; NID:g10581552; PIDN:AAG20275.1; GSPDB:G  
C:Genetics:  
A:Gene: ggt

Query Match 51.3%; Score 39; DB 2; Length 162;  
Best Local Similarity 53.3%; Pred. No. 34;  
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 KKKKAAAVLLPVLL 15  
DB 88 RRRRVFAAAVLLPVL 102

RESULT 13  
JC5847  
chitinase (EC 3.2.1.14) III C10923 - rice  
N:Alternate names: 1,4-beta-poly-N-acetylglucosaminidase; chitodextrinase; poly-beta-glu  
C:Species: Oryza sativa (rice)  
C:Date: 05-Mar-1998 #sequence\_revision 13-Mar-1998 #text\_change 09-Jul-2004  
A:Accession: JC5847  
R:Nagasaki, H.; Yamamoto, K.; Shomura, A.; Koga-Ban, Y.; Takasuga, A.; Yano, M.; Minobe,  
DNA Res. 4, 373-385, 1997  
A:Title: Rice class III chitinase homologues isolated by random cloning of rice cDNAs.  
A:Reference number: JC5841; MUID:98162724; PMID:9501993  
A:Accession: JC5847  
A:Status: nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 1-304 <NAG>  
A:CROSS-references: UNIPROT:Q9SXY2; UNIPROT:O49830; DBJ:D55714; NID:g2596232; PIDN:BA42  
C:Comment: This protein hydrolyzes poly-beta-1,4-N-acetyl glucosamine (chitin), and act  
C:Superfamily: plant chitinase III  
C:Keywords: glycosidase; hydrolase  
F:33-304/Product: chitinase III C10923 1 #status predicted <MAT>  
F:155/Active site: Asp #status predicted

Query Match 51.3%; Score 39; DB 2; Length 304;  
Best Local Similarity 50.0%; Pred. No. 61;  
Matches 8; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 1 KKKKAAAVLLPVLL 16  
DB 4 QRRRSATVLLSLL 19

RESULT 14  
S78242  
cytochrome c-type synthesis protein homolog - Odontella sinensis chloroplast  
N:Alternate names: hypothetical protein 312  
C:Species: chloroplast Odontella sinensis  
C:Date: 17-Feb-1998 #sequence\_revision 26-Feb-1998 #text\_change 09-Jul-2004  
A:Accession: S78242  
R:Kowallik, K.V.; Stoebe, B.; Schaffran, I.; Kroth-Pancic, P.; Freier, U.  
Plant Mol. Biol. Rep. 13, 336-342, 1995  
A:Title: The Chloroplast Genome of a chlorophyll a+c- containing Alga, Odontella sinensis  
A:Reference number: S78238  
A:Accession: S78242  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-312 <KOW>  
A:CROSS-references: UNIPROT:P49523; EMBL:Z67753; NID:g1185127; PIDN:CAA31615.1; PID:g118  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1995  
C:Genetics:  
A:Gene: ycf5  
A:Genome: chloroplast  
A:Superfamily: cytochrome c-type synthesis protein  
C:Keywords: chloroplast

Query Match 51.3%; Score 39; DB 2; Length 312;  
Best Local Similarity 50.0%; Pred. No. 63;  
Matches 8; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 KKKKAAAVLLPVLL 16  
DB 93 KTKSKLVGALLIPVAL 108

RESULT 15

D87506  
glycosyl transferase family protein CC2077 [imported] - Caulobacter crescentus  
C:Species: Caulobacter crescentus  
C>Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 09-Jul-2004  
C:Accession: D87506  
R:Niekman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.  
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon  
n, J.; Ermlaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.  
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001  
A:Title: Complete Genome Sequence of Caulobacter crescentus.  
A:Reference number: A87249; MUID:21173698; PMID:11259647  
A:Accession: D87506  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-318 <STO>  
A:Cross-references: UNIPROT:Q9A6L6; GB:AE005673; NID:gl3423558; PIDN:AAK24048.1; GSPDB:G  
C:Genetics:  
A:Gene: CC2077

Query Match 51.3%; Score 39; DB 2; Length 318;  
Best Local Similarity 43.8%; Pred. No. 64;  
Matches 7; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 2 KKRKAAAVLLPVLLA 17  
:|||||:|:|:|:|:  
Db 276 RRRKALALLAPILLS 291

RESULT 16  
T42653  
hypothetical protein DKFZp434F2015.1 - human (fragment)  
C:Species: Homo sapiens (man)  
C>Date: 11-Jan-2000 #sequence\_revision 11-Jan-2000 #text\_change 09-Jul-2004  
C:Accession: T42653  
R:Blöcker, H.; Boecher, M.; Brandt, P.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.  
submitted to the Protein Sequence Database, November 1999  
A:Reference number: Z22230  
A:Accession: T42653  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-437 <AAA>  
A:Cross-references: UNIPROT:Q9UFC5; EMBL:AL133050  
A:Experimental source: adult testis; clone DKFZp434F2015  
C:Genetics:  
A:Note: DKFZp434F2015.1

Query Match 51.3%; Score 39; DB 2; Length 437;  
Best Local Similarity 66.7%; Pred. No. 86;  
Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 KKRKAAAVLL 12  
:|||||:|:|:|:|:  
Db 424 KKRKRAAALVL 435

RESULT 17  
A70628  
probable acid-CoA ligase (EC 6.2.1.-) fadD2 - Mycobacterium tuberculosis (strain H37Rv)  
C:Species: Mycobacterium tuberculosis  
C>Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 09-Jul-2004  
C:Accession: A70628  
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.  
; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Hoolroyd, S.  
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.  
Nature 393, 537-544, 1998  
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome  
A:Reference number: A70500; MUID:98295987; PMID:9634230  
A:Accession: A70628  
A>Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-560 <COL>  
A:Cross-references: UNIPROT:P95227; GB:Z86089; GB:AL123456; NID:g3261711; PIDN:CAB06682.

A:Experimental source: strain H37Rv  
C:Genetics:  
A:Gene: fadD2  
C:Superfamily: 4-coumarate-CoA ligase; acetate-CoA ligase homology  
C:Keywords: acid-thiol ligase  
F:107-557/Domain: acetate-CoA ligase homology <ACL>  
Query Match 51.3%; Score 39; DB 2; Length 560;  
Best Local Similarity 43.8%; Pred. No. 1.1e+02;  
Matches 7; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 2 KKRKAAAVLLPVLLA 17  
:|||||:|:|:|:|:  
Db 304 EXKATSMVVPVMLS 319

RESULT 18  
S53978  
PSI protein - yeast (Saccharomyces cerevisiae)  
N:Alternate names: protein YM952.10c; protein YMR308c  
C:Species: Saccharomyces cerevisiae  
C>Date: 08-Jul-1995 #sequence\_revision 01-Sep-1995 #text\_change 09-Jul-2004  
C:Accession: S53978; S20189  
R:Connor, R.; Churcher, C.M.  
submitted to the EMBL Data Library, April 1995  
A:Reference number: S53969  
A:Accession: S53978  
A:Molecule type: DNA  
A:Residues: 1-1089 <CON>  
A:Cross-references: UNIPROT:P32337; EMBL:Z49212; NID:g798940; PID:g798950; MIPS:YMR308c  
R:Chow, T.Y.K.; Ash, J.; Dignard, D.; Thomas, D.Y.  
submitted to the EMBL Data Library, December 1991  
A:Description: Screening and identification of a gene, PSE-1, that affects protein sec  
A:Reference number: S20189  
A:Accession: S20189  
A:Molecule type: DNA  
A:Residues: 1-64, 'S', '66-1076, 'A', '1078-1089 <CHO>  
A:Cross-references: EMBL:Z11538; NID:g4245; PID:g4246  
C:Genetics:  
A:Gene: SGD:PSE1  
A:Cross-references: SGD:S0004925; MIPS:YMR308c  
A:Map position: 13R  
C:Keywords: transmembrane protein  
F:401-417/Domain: transmembrane #status predicted <TM1>  
F:616-632/Domain: transmembrane #status predicted <TM2>  
F:719-735/Domain: transmembrane #status predicted <TM3>  
F:874-890/Domain: transmembrane #status predicted <TM4>

Query Match 51.3%; Score 39; DB 2; Length 1089;  
Best Local Similarity 53.8%; Pred. No. 2e+02;  
Matches 7; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 5 KRAAAVLLPVLLA 17  
:|||||:|:|:|:|:  
Db 718 RAAGATLIPILLS 730

RESULT 19  
T48162  
hypothetical protein T1008.110 - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C>Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 09-Jul-2004  
C:Accession: T48162  
R:Bevan, M.; Pohl, T.; Weizenecker, T.; Bancroft, I.; Mewes, H.W.; Lemcke, K.; Mayer, K.  
submitted to the Protein Sequence Database, March 2000  
A:Reference number: Z24486  
A:Accession: T48162  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-1467 <BEV>  
A:Cross-references: UNIPROT:Q9M033; EMBL:AL161746  
A:Experimental source: cultivar Columbia; BAC clone T1008  
C:Genetics:



RESULT 21  
C82807  
30S ribosomal protein S21 XF04334 [imported] - Xylella fastidiosa (strain 9a5c)  
C:Species: Xylella fastidiosa  
C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 09-Jul-2004  
C:Accession: C82807  
R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequencing  
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.  
A:Reference number: A83515; MUID:20355717; PMID:10910347  
A:Note: for a complete list of authors see reference number A59328 below  
A:Accession: C82807  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 171 <STM>  
A:Cross-references: UNIPROT:Q9PG68; GB:AE003894; GB:AE003849; NID:g9105267; PIDN:AAF8324  
A:Experimental source: strain 9a5c  
R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; Brites, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carreir, H. submitted to Genbank, June 2000  
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laig Chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins E.A.; Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.; F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A.; Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasaki A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silva

QY 2 KKKKAAAAVLLPVLL 16  
: : ||||| :  
Db 5 RRRRLAAVLLSVLL 19

## RESULT 24

T05461  
hypothetical protein F7K2.250 - Arabidopsis thaliana (fragment)  
C:Species: Arabidopsis thaliana (mouse-ear cross)  
C:Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 23-Jul-1999  
C:Accession: T05461  
R:Bevan, M.; Wedler, H.; Wambutt, R.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.X.; Schuelle  
submitted to the Protein Sequence Database, November 1998  
A:Reference number: Z15416  
A:Accession: T05461  
A:Molecule type: DNA  
A:Residues: 1-287 <BEV>  
A:Cross-references: EMBL:AL033545  
A:Experimental source: cultivar Columbia; BAC clone F7K2  
C:Genetics:  
A:Map position: 4  
A:Introns: 113/1; 146/3; 212/3; 269/2  
A:Note: F7K2.250

Query Match 50.0%; Score 38; DB 2; Length 287;  
Best Local Similarity 60.0%; Pred. No. 85;  
Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 KKKKAAAAVLLPVLL 15  
: ||||| :  
Db 135 EKKKEBAVAMGVL 149

## RESULT 25

G75619  
probable hemin ABC transporter, permease protein - Deinococcus radiodurans (strain R1)  
C:Species: Deinococcus radiodurans  
C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004  
C:Accession: G75619  
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;  
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma  
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.  
Science 286, 1571-1577, 1999  
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.  
A:Reference number: A75250; MUID:20036896; PMID:10567266  
A:Accession: G75619  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-354 <WHI>  
A:Cross-references: UNIPROT:Q9RZUG; GB:AE001826; NID:G6460827; PIDN:AAF12576.1; PID:G646  
A:Experimental source: strain R1  
C:Genetics:  
A:Gene: DRB0015  
A:Map position: megaplasmid  
A:Genome: plasmid  
A:Note: plasmid MP1  
C:Superfamily: vitamin B12 transport protein btuc

Query Match 50.0%; Score 38; DB 2; Length 354;  
Best Local Similarity 50.0%; Pred. No. 1e+02;  
Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 KKKKAAAAVLLPVLL 16  
: : ||||| :  
Db 14 QRSRAALSLVLLPLL 29

## RESULT 26

H75376  
hypothetical protein - Deinococcus radiodurans (strain R1)  
C:Species: Deinococcus radiodurans  
C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004  
C:Accession: H75376

R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;  
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma  
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.  
Science 286, 1571-1577, 1999  
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.  
A:Reference number: A75250; MUID:20036896; PMID:10567266  
A:Accession: H75376  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-411 <WHI>  
A:Cross-references: UNIPROT:Q9RU03; GB:AE002003; GB:AE000513; NID:G6459358; PIDN:AAF111  
A:Experimental source: strain R1  
C:Genetics:  
A:Gene: DR1594  
A:Map position: 1

Query Match 50.0%; Score 38; DB 2; Length 411;  
Best Local Similarity 60.0%; Pred. No. 1.2e+02;  
Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 3 KKKAAAAVLLPVLLA 17  
: ||||| :  
Db 62 KRPLFAALLCPVLLA 76

## RESULT 27

G87227  
acyl-CoA synthase [imported] - Mycobacterium leprae  
C:Species: Mycobacterium leprae  
C:Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 09-Jul-2004  
C:Accession: G87227  
R:Coile, S.T.; Eiglmeier, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.; H  
R.; Davies, R.M.; Devlin, K.; Duthoy, S.; Felwell, T.; Fraser, A.; Hamlin, N.; Holroyd  
eam, M.A.; Rutherford, K.M.  
Nature 409, 1007-1011, 2001  
A:Authors: Rutter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squares, R.; S  
A:Title: Massive gene decay in the leprosy bacillus.  
A:Reference number: A86909; MUID:21128732; PMID:11234002  
A:Accession: G87227  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-548 <STO>  
A:Cross-references: UNIPROT:Q9CD27; GB:AL450380; NID:gl3093799; PIDN:CAC32077.1; GSPDB:(  
A:Gene: fadD2  
C:Superfamily: 4-coumarate-CoA ligase; acetate-CoA ligase homology

Query Match 50.0%; Score 38; DB 2; Length 548;  
Best Local Similarity 43.8%; Pred. No. 1.6e+02;  
Matches 7; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 2 KKKKAAAAVLLPVLLA 17  
: : ||||| :  
Db 289 EKHQATAMVVVFMVLS 304

## RESULT 28

T04562  
hypothetical protein T12H17.60 - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cross)  
C:Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 09-Jul-2004  
C:Accession: T04562  
R:Bevan, M.; Hilbert, H.; Braun, M.; Holzer, E.; Brandt, A.; Duesterhoeft, A.; Bancroft  
submitted to the Protein Sequence Database, February 1998  
A:Reference number: Z15377  
A:Accession: T04562  
A:Molecule type: DNA  
A:Residues: 1-627 <BEV>  
A:Cross-references: UNIPROT:O49648; EMBL:AL021635  
A:Experimental source: cultivar Columbia; BAC clone T12H17  
C:Genetics:  
A:Map position: 4  
A:Introns: 12/1; 111/1; 144/3; 272/2; 290/3; 342/3; 399/1; 420/3; 467/3; 506/3; 531/1;

A>Note: T12H17.60

Query Match 50.0%; Score 38; DB 2; Length 627;  
Best Local Similarity 60.0%; Pred. No. 1.8e+02;  
Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 1 KKKRKAARAAVLPVL 15  
Db 133 EKRRKEAAVANGVL 147

RESULT 29

T46560 swi2 protein - fission yeast (Schizosaccharomyces pombe)

C:Species: Schizosaccharomyces pombe  
C:Date: 18-Feb-2000 #sequence\_revision 18-Feb-2000 #text\_change 09-Jul-2004  
C:Accession: T46560 T37852

R:Schmidt, H.  
submitted to the EMBL Data Library, November 1995

A:Reference number: Z23067

A:Accession: T46560

A>Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-722 <SCH>

A:Cross-references: UNIPROT:Q10668; EMBL:X93590; PIDN:CAA63788.1

A:Experimental source: strain L972

R:Murphy, L.; Harris, D.; Barrell, B.G.; Rajandream, M.A.; Wood, V.

submitted to the EMBL Data Library, September 1997

A:Reference number: Z21749

A:Accession: T37852

A:Molecule type: DNA

A:Residues: 178-722 <MUR>

A:Cross-references: EMBL:Z99162; PIDN:CAB16231.1; GSPDB:GN00066; SPDB:SPAC17G6.20C

C:Genetics:

A:Gene: swi2

A:Map position: 1

Query Match 50.0%; Score 38; DB 2; Length 722;  
Best Local Similarity 50.0%; Pred. No. 2e+02;  
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 KKKRKAARAAVLPVLL 16  
Db 265 KKPRKVFDAVPIPTI 280

RESULT 30

F85047

hypothetical protein AT4G03760 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 09-Jul-2004

C:Accession: F85047

R:anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring

Nature 402, 769-777, 1999

A:Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.

A:Reference number: A85001; MUID:20083488; PMID:10617198

A:Accession: F85047

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-723 <STO>

A:Cross-references: UNIPROT:Q9SV54; GB:NC\_001268; NID:g7270699; PIDN:CAB77861.1; GSPDB:G

C:Genetics:

A:Gene: AT4G03760

A:Map position: 4

Query Match 50.0%; Score 38; DB 2; Length 723;  
Best Local Similarity 53.8%; Pred. No. 2e+02;  
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 KKKRKAARAAVLP 13  
Db 138 KKKRKKKKVIMP 150

RESULT 31

T20122

hypothetical protein F25B3.1 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004

C:Accession: T20122; T21324

R:McMurray, A.

submitted to the EMBL Data Library, April 1996

A:Reference number: Z19225

A:Accession: T20122

A>Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-901 <WIL>

A:Cross-references: UNIPROT:Q18749; EMBL:Z70750; PIDN:CAA94750.1; GSPDB:GN00023; CESP:F25B3.1

A:Experimental source: clone C50F4

R:Gardner, A.

submitted to the EMBL Data Library, April 1996

A:Reference number: Z19406

A:Accession: T21324

A>Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-901 <WIL>

A:Cross-references: EMBL:Z70752; PIDN:CAA94758.1; GSPDB:GN00023; CESP:F25B3.1

A:Experimental source: clone F25B3

C:Genetics:

A:Gene: CESP:F25B3.1

A:Map position: 5

A:Introns: 25/3; 47/2; 199/1; 242/1; 315/2; 544/3; 600/1; 626/1; 646/1; 666/3; 706/3; 706/3; 75/3

Query Match 50.0%; Score 38; DB 2; Length 901;

Best Local Similarity 68.8%; Pred. No. 2.5e+02;

Matches 11; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 KKKRKAARAAVLPVLL 16  
Db 108 KKKRKAATAAVPLNRL 123

RESULT 32

T00057

hypothetical protein KIAA0423 - human (fragment)

C:Species: Homo sapiens (man)

C:Date: 22-Jan-1999 #sequence\_revision 22-Jan-1999 #text\_change 09-Jul-2004

C:Accession: T00057

R:Ishikawa, K.; Nagase, T.; Nakajima, D.; Seki, N.; Ohira, M.; Miyajima, N.; Tanaka, A.,

submitted to the EMBL Data Library, October 1997

A:Description: Prediction of the coding sequences of unidentified human genes. VIII. The

A:Reference number: Z14082

A:Accession: T00057

A>Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: mRNA

A:Residues: 1-1696 <ISH>

A:Cross-references: UNIPROT:Q9Y4F4; EMBL:AB007883; NID:d1179759; PIDN:BAA24853.1; PID:d1

A:Experimental source: brain

C:Genetics:

A>Note: KIAA0423

Query Match 50.0%; Score 38; DB 2; Length 1696;

Best Local Similarity 58.3%; Pred. No. 4.5e+02;

Matches 7; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 5 KAAAVALLPVLL 16  
Db 207 RASTALLPILL 218

RESULT 33

AI1291

glycerol dehydrogenase homolog lmo1737 [imported] - Listeria monocytogenes (strain EGD-e

C:Species: Listeria monocytogenes

C:Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 16-Aug-2004

C:Accession: AI1291

R.; Glaeser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; Science 294, 849-852, 2001

A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Maok, C.; Schluteter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend, A.; Title: Comparative genomics of *Listeria* species.

A:Reference number: AB1077; MUID:21537279; PMID:11679669

A:Accession: A1291

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-368 <GLA>

A:Cross-references: UNIPROT:Q8V6R0; GB:NC\_003210; PIDN:CAC99815.1; PID:g16411191; GSPDB:A:Experimental source: strain EGB-e

C:Genetics:

C:Superfamily: Glycerol dehydrogenase; lactaldehyde reductase homology

Query Match 49.3%; Score 37.5; DB 2; Length 368;

Best Local Similarity 76.9%; Pred. No. 1.3e+02;

Matches 10; Conservative 2; Mismatches 0; Indels 1; Gaps 1;

QY 5 KAAAVAL-LPVL 16

DB 103 KAAAVALRLPVII 115

RESULT 34

AG1663

glycerol dehydrogenase homolog lin1848 [imported] - *Listeria innocua* (strain Clip11262)

C:Species: *Listeria innocua*

C:Date: 27-Nov-2001 #sequence\_revision 27-Nov-2001 #text\_change 16-Aug-2004

C:Accession: AG1663

R.; Glaeser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; Science 294, 849-852, 2001

A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Maok, C.; Schluteter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlend, D.; Jones, L.M.; Karst, U.

A:Title: Comparative genomics of *Listeria* species.

A:Reference number: AB1077; MUID:21537279; PMID:11679669

A:Accession: AG1663

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-368 <GLA>

A:Cross-references: UNIPROT:Q92AS1; GB:AL592022; PIDN:CAC97079.1; PID:g16414350; GSPDB:A:Experimental source: strain Clip11262

C:Genetics:

C:Superfamily: Glycerol dehydrogenase; lactaldehyde reductase homology

Query Match 49.3%; Score 37.5; DB 2; Length 368;

Best Local Similarity 76.9%; Pred. No. 1.3e+02;

Matches 10; Conservative 2; Mismatches 0; Indels 1; Gaps 1;

QY 5 KAAAVAL-LPVL 16

DB 103 KAAAVALRLPVII 115

RESULT 35

D84231

hypothetical protein Vng0737h [imported] - *Halobacterium* sp. NRC-1

C:Species: *Halobacterium* sp. NRC-1

C:Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 09-Jul-2004

C:Accession: D84231

R.; Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Laaky, S.; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jabl, Jung, K.H.; Alam, M.; Freitas, T.

Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000

A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li

A:Title: Genome sequence of *Halobacterium* species NRC-1.

A:Reference number: A84160; MUID:20504483; PMID:11016950

A:Accession: D84231

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-74 <STO>

A:Cross-references: UNIPROT:Q9HRE2; GB:AE004437; NID:g10580318; PIDN:AAG19216.1; GSPDB:A:Genetics:

A:Gene: VNG0737H

Query Match 48.7%; Score 37; DB 2; Length 74;

Best Local Similarity 61.5%; Pred. No. 36;

Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 4 KAAAVAL-LPVL 16

DB 14 RKQAAVALPFL 26

RESULT 36

E70048

hypothetical protein yvtA - *Bacillus subtilis*

C:Species: *Bacillus subtilis*

C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 28-Jul-2000

C:Accession: E70048

R.; Kunz, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter, C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch, A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.; Nature 390, 249-256, 1997

A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galle, iech, J.; Harwood, C.R.; Hensaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F. Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois, A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maues, Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetell, Rieger, M.; Rivolca, C.; Rocha, E.; Roche, B.; Rose, M.; Sadate, Y.; Sato, T.; Scanlon, A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Sero, akeuchi, M.; Tanakoshi, A.; Tanaka, T.; Terpsira, P.; Tognoni, A.; Tosato, V.; Uchiyama, T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, A:Authors: Yoshikawa, H.P.; Zumstein, E.; Yoshikawa, H.; Danchin, A.

A:Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.

A:Reference number: A95980; MUID:19804033; PMID:9384377

A:Accession: E70048

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-121 <KUN>

A:Cross-references: GB:Z99120; GB:AL009126; NID:g2635613; PIDN:CAB15230.1; PID:el184379, A:Experimental source: strain 168

C:Genetics:

A:Gene: yvtA

C:Superfamily: *Bacillus subtilis* hypothetical protein yvtA

Query Match 48.7%; Score 37; DB 2; Length 121;

Best Local Similarity 53.3%; Pred. No. 56;

Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 KKKRAAAVALPVL 15

DB 65 KKKRAAAVALPVL 79

RESULT 37

A87461

hypothetical protein CCl709 [imported] - *Caulobacter crescentus*

C:Species: *Caulobacter crescentus*

C:Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 09-Jul-2004

C:Accession: A87461

R.; Nierman, W.C.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon, J.; Esmolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M. n, J.; Laub, M.T.; DeBoy, R.T.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J. n, J.; Esmolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M. Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001

A:Title: Complete genome sequence of *Caulobacter crescentus*.

A:Reference number: A87249; MUID:21173698; PMID:11259647

A:Accession: A87461

A>Status: preliminary

A:Molecule type: DNA

COMOZ, R.; DAVAZ, R.; DEZZI, R.; FORTI, R.; GENTILE, S.; HAMANN, N.; MORFORD, J.

A:Molecule type: DNA  
A:Residues: 1-412 <KUR>  
A:Cross-references: UNIPROT:Q8U961; GB:AE007870; PID:AAK89551.1; PID:g15159433; GSPDB:C

**Genetics:**



```
A:Residues: 1-607 <TOM>
A:Cross-references: UNIPROT:O25379; GB:AE000580; GB:AE000511; NID:G2313792; PIDN:AAD0773
C:Genetics:
A:Start Codon: GTG
C:Keywords: ATP; nucleotide binding; P-loop
F:124-131/Region: nucleotide-binding motif A (P-loop)
F:248-253/Region: nucleotide-binding motif B
F:252-255/Region: DEAH motif

Query Match      48.7%; Score 37; DB 2; Length 607;
Best Local Similarity 43.8%; Pred. No. 2.5e+02;
Matches 7; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 KKKRKAANAALVLPVLL 16
   |||||:|:|:|
Db 535 KAKRKRKGVILPILAL 550

RESULT 47
T05209
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C:Accession: T05209
R:Bevan, M.; Vitale, D.; Liguori, R.; Argiriou, A.; De Simone, V.; Bancroft, I.; Mewes,
submitted to the Protein Sequence Database, February 1998
A:Reference number: Z15403
A:Accession: T05209
A:Molecule type: DNA
A:Residues: 1-939 <BEV>
A:Cross-references: UNIPROT:O49467; EMBL:AL021768; GSPDB:GN00062; ATSP:F24J7.50
A:Experimental source: cultivar Columbia; BAC clone F24J7
C:Genetics:
A:Gene: ATSP:F24J7.50
A:Map position: 4
A:Introns: 278/3; 299/3; 326/2; 336/2; 376/3; 401/3; 453/1; 492/3; 578/2; 635/3; 663/3;

Query Match      48.7%; Score 37; DB 2; Length 939;
Best Local Similarity 52.9%; Pred. No. 3.8e+02;
Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 KKKRKAANAALVLPVLLA 17
   |||||:|:|:|
Db 420 KNAIKKAVADLLPILVA 436

RESULT 48
T30877
adenylate cyclase (EC 4.6.1.1) - Trypanosoma cruzi
C:Species: Trypanosoma cruzi
C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 09-Jul-2004
C:Accession: T30877
R:Taylor, M.C.
submitted to the EMBL Data Library, July 1999
A:Reference number: Z20916
A:Accession: T30877
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1248 <TAY>
A:Cross-references: UNIPROT:O96762; EMBL:AJ012096; PIDN:CAA09919.2
C:Genetics:
A:Gene: adcl
C:Superfamily: Trypanosoma adenylate cyclase; yeast adenylate cyclase catalytic domain H
C:Keywords: phosphorus-oxygen lyase

Query Match      48.7%; Score 37; DB 2; Length 1248;
Best Local Similarity 50.0%; Pred. No. 4.9e+02;
Matches 7; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 3 KKKRKAANAALVLPVLL 16
   |||||:|:|:|
Db 257 RRTAGAYILLPLIL 270
```

```
RESULT 49
E84601
probable retroelement pol polyprotein [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C:Accession: E84601
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter,
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: E84601
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1333 <STO>
A:Cross-references: UNIPROT:Q9SJT2; GB:AE002093; NID:G4567277; PIDN:AAD23690.1; GSPDB:G
C:Genetics:
A:Gene: At2G21460
A:Map position: 2
C:Superfamily: retrovirus-related polyprotein

Query Match      48.7%; Score 37; DB 2; Length 1333;
Best Local Similarity 50.0%; Pred. No. 5.2e+02;
Matches 7; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 1 KKKRKAANAALVLPV 14
   |||||:|:|:|
Db 69 EKKRKARSAIVLSV 82

RESULT 50
F84486
probable retroelement pol polyprotein [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C:Accession: F84486
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter,
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: F84486
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1356 <STO>
A:Cross-references: UNIPROT:Q9SH77; GB:AE002093; NID:G4895171; PIDN:AAD32759.1; GSPDB:G
C:Genetics:
A:Gene: At2G07550
A:Map position: 2
C:Superfamily: retrovirus-related polyprotein

Query Match      48.7%; Score 37; DB 2; Length 1356;
Best Local Similarity 50.0%; Pred. No. 5.3e+02;
Matches 7; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 1 KKKRKAANAALVLPV 14
   |||||:|:|:|
Db 69 EKKRKARSAIVLSV 82

Search completed: June 2, 2005, 01:35:52
Job time : 19.2812 secs
```

THIS PAGE BLANK (USPTO)



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 2, 2005, 01:21:36 ; Search time 124.312 Seconds  
(without alignments)  
52.890 Million cell updates/sec

Title: US-10-764-235-24

Perfect score: 76

Sequence: 1 KKKRKAARVLLPVLIA 17

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 750 summaries

Database :

A\_Geneseq\_16Dec04:\*

1: Geneseqp1980s:\*

2: Geneseqp1990s:\*

3: Geneseqp2000s:\*

4: Geneseqp2001s:\*

5: Geneseqp2002s:\*

6: Geneseqp2003as:\*

7: Geneseqp2003bs:\*

8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	76	100.0	17	4	Aau00654 Human mem
2	76	100.0	17	7	Adb88797 Membrane
3	76	100.0	17	8	Abg75448 Membrane
4	53.5	70.4	29	3	Aay55835 SV4OMEM p
5	53.5	70.4	29	6	Abu69594 Human NP-
6	53.5	70.4	30	2	Aaw56391 SV4OMEM p
7	53.5	70.4	30	3	Aay55812 SV4OMEM p
8	53	69.7	16	4	Aau00633 Human mem
9	53	69.7	16	6	Abp70215 Amino aci
10	53	69.7	16	7	Adb88776 Membrane
11	53	69.7	20	6	Abp70216 Amino aci
12	51	67.1	15	6	Abp70234 Amino aci
13	51	67.1	15	6	Abp70214 Amino aci
14	51	67.1	15	6	Abp70233 Amino aci
15	51	67.1	15	8	Abg75426 Membrane
16	51	67.1	19	4	Aau00655 Human tar
17	51	67.1	19	4	Aau00634 Human mem
18	51	67.1	19	7	Adb88777 Membrane
19	51	67.1	19	8	Abg75428 Membrane
20	51	67.1	21	4	Aau00632 Human mem
21	51	67.1	21	7	Adb88775 Membrane
22	51	67.1	22	3	Aay67959 Karp0si s
23	50	65.8	21	8	Adk50919 Human C35
24	50	65.8	33	8	Adk50920 Human C35
25	48	63.2	433	8	Adp54374 Human PRO

26	48	63.2	433	8	Adp23402 PRO poly p
27	47	61.8	14	4	Aau00637 Human mem
28	47	61.8	14	7	Adb88780 Membrane
29	47	61.8	14	8	Abg75427 Membrane
30	47	61.8	14	8	Abg75431 Membrane
31	47	61.8	16	4	Aau00635 Human mem
32	47	61.8	16	7	Adb88778 Membrane
33	47	61.8	16	8	Abg75429 Membrane
34	47	61.8	18	8	Adr40510 Peptide c
35	47	61.8	18	8	Adr40509 Peptide c
36	47	61.8	21	8	Adk50807 Human C35
37	47	61.8	21	8	Adk50911 Human C35
38	47	61.8	21	8	Adk50895 Human C35
39	47	61.8	21	8	Adk50915 Human C35
40	47	61.8	21	8	Adk50899 Human C35
41	47	61.8	22	8	Adk50923 Human C35
42	47	61.8	26	2	Aaw56392 HIV-1MEM
43	47	61.8	26	3	Aay55813 HIV-1MEM
44	47	61.8	33	8	Adk50900 Human C35
45	47	61.8	33	8	Adk50808 Human C35
46	47	61.8	33	8	Adk50916 Human C35
47	47	61.8	33	8	Adk50896 Human C35
48	47	61.8	33	8	Adk50912 Human C35
49	47	61.8	34	8	Adk50924 Human C35
50	46.5	61.2	29	2	Aaw56414 Nuclear t
51	46	60.5	15	6	Abp70217 Amino aci
52	46	60.5	20	3	Aay67963 Karp0si s
53	46	60.5	21	8	Adk50835 Human C35
54	46	60.5	33	8	Adk50836 Human C35
55	46	60.5	365	6	Abu19600 Protein e
56	45	59.2	21	8	Adk50831 Human C35
57	45	59.2	21	8	Adk50891 Human C35
58	45	59.2	33	8	Adk50892 Human C35
59	45	59.2	33	8	Adk50832 Human C35
60	44	57.9	19	2	Aay44170 Fragment
61	44	57.9	21	8	Adk50903 Human C35
62	44	57.9	22	8	Adk50907 Human C35
63	44	57.9	25	2	Aay44169 Fragment
64	44	57.9	33	8	Adk50904 Human C35
65	44	57.9	34	8	Adk50908 Human C35
66	44	57.9	802	8	Adi45285 Rice isop
67	43	56.6	10	2	Aay44163 Truncated
68	43	56.6	11	2	Aay44164 Truncated
69	43	56.6	11	4	Aau00645 Human mem
70	43	56.6	11	7	Adb88788 Membrane
71	43	56.6	11	8	Abg75439 Membrane
72	43	56.6	12	2	Aay44160 Membrane-
73	43	56.6	12	4	Aae02980 Hydrophob
74	43	56.6	12	4	Aau00644 Human mem
75	43	56.6	12	4	Aau00646 Human mem
76	43	56.6	12	5	Abg78983 Cell pene
77	43	56.6	12	5	Au78350 Grb2 SH2
78	43	56.6	12	5	Au78350 Grb2 SH2
79	43	56.6	12	5	Abb81929 Cystic fi
80	43	56.6	12	5	Abg75508 Signal-se
81	43	56.6	12	5	Abb81178 Grb2 SH2
82	43	56.6	12	5	Aau77231 Synthetic
83	43	56.6	12	6	Aae32065 Transport
84	43	56.6	12	7	Adb88789 Membrane
85	43	56.6	12	7	Adc22455 Protein-d
86	43	56.6	12	7	Adg28018 Kaposi' s
87	43	56.6	12	7	Adh76185 Transduct
88	43	56.6	12	7	Adl88654 MPS (Kapo
89	43	56.6	12	7	Adn60178 Novel rec
90	43	56.6	12	7	Adg12916 Cytoplaem
91	43	56.6	12	8	Abg75425 Kaposi' s
92	43	56.6	12	8	Abg75438 Membrane
93	43	56.6	12	8	Abg75440 Membrane
94	43	56.6	12	8	Adj78876 Src homol
95	43	56.6	12	8	Adl14687 Cardiant
96	43	56.6	12	8	Adk15575 Membrane
97	43	56.6	12	8	Ado26467 Grb2 SH2
98	43	56.6	12	8	Adq60180 Mammalian

99	43	56.6	12	8	ADQ91936	Adq91936 Kaposi fi	172	41	53.9	549	8	ADP29329	Adp29329 Human sec
100	43	56.6	12	8	ADQ80671	Adq80671 Kaposi FG	173	41	53.9	579	2	AAR86406	Aar86406 Human mat
101	43	56.6	12	8	ADRS1973	Adrs1973 Heat shoc	174	41	53.9	582	2	AAR86407	Aar86407 Human mat
102	43	56.6	12	8	ADRS17602	Adrs17602 Membrane	175	41	53.9	582	2	AAR75648	Aar75648 Human pla
103	43	56.6	13	4	AAU006636	Aau006636 Human mem	176	41	53.9	582	2	AAW52134	Aaw52134 Rabbit me
104	43	56.6	13	5	ABG68407	Abg68407 Membrane	177	41	53.9	582	2	AAW52134	Aaw52134 Rabbit me
105	43	56.6	13	7	ADB88779	Adb88779 Membrane	178	41	53.9	582	2	AAW52134	Aaw52134 Rabbit me
106	43	56.6	13	8	ABG75430	Abg75430 Membrane	179	41	53.9	582	4	AAE10423	Aae10423 Human mat
107	43	56.6	17	4	AAU006653	Aau006653 Human mem	180	41	53.9	582	5	AAU84294	Aau84294 Human end
108	43	56.6	17	7	ADB88796	Adb88796 Membrane	181	41	53.9	582	5	AAE21037	Aae21037 Human mem
109	43	56.6	17	7	ABG75447	Abg75447 Membrane	182	41	53.9	582	5	AAW50865	Aaw50865 Matrix me
110	43	56.6	17	8	ADRS17601	Adrs17601 Membrane	183	41	53.9	582	7	ADC15498	Adc15498 Human bas
111	43	56.6	17	8	ADR88404	Adr88404 Mouse MTS	184	41	53.9	582	7	ADG64179	Adg64179 Human Pro
112	43	56.6	20	5	ABG79150	Abg79150 Cell pene	185	41	53.9	582	7	ADG64177	Adg64177 Rat Prote
113	43	56.6	21	5	ABG79149	Abg79149 Cell pene	186	41	53.9	582	7	ADP13708	Adp13708 Tumor-Ass
114	43	56.6	21	8	ADKS50791	Adks50791 Human C35	187	41	53.9	582	7	ADN95889	Adn95889 Human BEC
115	43	56.6	21	8	ADKS50887	Adks50887 Human C35	188	41	53.9	582	8	ADN07703	Adn07703 Human mat
116	43	56.6	21	8	ADKS50819	Adks50819 Human C35	189	41	53.9	582	8	ABM81541	Abm81541 Tumor-as
117	43	56.6	21	8	ADKS50843	Adks50843 Human C35	190	40.5	53.3	582	8	ADP23426	Adp23426 PRO polyP
118	43	56.6	21	8	ADKS50795	Adks50795 Human C35	191	40.5	53.3	582	8	ABM81541	Abm81541 Tumor-as
119	43	56.6	21	8	ADKS50815	Adks50815 Human C35	192	40.5	53.3	582	8	ABM81541	Abm81541 Tumor-as
120	43	56.6	21	8	ADKS50823	Adks50823 Human C35	193	40	52.6	582	8	ABM81541	Abm81541 Tumor-as
121	43	56.6	21	8	ADKS50803	Adks50803 Human C35	194	40	52.6	582	8	ABM81541	Abm81541 Tumor-as
122	43	56.6	21	8	ADKS50811	Adks50811 Human C35	195	40	52.6	582	8	ABM81541	Abm81541 Tumor-as
123	43	56.6	21	8	ADKS50799	Adks50799 Human C35	196	40	52.6	582	8	ABM81541	Abm81541 Tumor-as
124	43	56.6	22	8	ADKS50839	Adks50839 Human C35	197	40	52.6	582	8	ABM81541	Abm81541 Tumor-as
125	43	56.6	25	8	ADKS50847	Adks50847 Human C35	198	40	52.6	582	8	ABM81541	Abm81541 Tumor-as
126	43	56.6	26	8	ADKS50883	Adks50883 Human C35	199	40	52.6	582	8	ABM81541	Abm81541 Tumor-as
127	43	56.6	27	8	ADKS50879	Adks50879 Human C35	200	40	52.6	582	8	ABM81541	Abm81541 Tumor-as
128	43	56.6	27	8	ADKS50827	Adks50827 Human C35	201	40	52.6	582	8	ABM81541	Abm81541 Tumor-as
129	43	56.6	33	8	ADKS50804	Adks50804 Human C35	202	40	52.6	582	8	ABM81541	Abm81541 Tumor-as
130	43	56.6	33	8	ADKS50851	Adks50851 Human C35	203	39	51.3	582	8	ABM81541	Abm81541 Tumor-as
131	43	56.6	33	8	ADKS50792	Adks50792 Human C35	204	39	51.3	582	8	ABM81541	Abm81541 Tumor-as
132	43	56.6	33	8	ADKS50844	Adks50844 Human C35	205	39	51.3	582	8	ABM81541	Abm81541 Tumor-as
133	43	56.6	33	8	ADKS50820	Adks50820 Human C35	206	39	51.3	582	8	ABM81541	Abm81541 Tumor-as
134	43	56.6	33	8	ADKS50796	Adks50796 Human C35	207	39	51.3	582	8	ABM81541	Abm81541 Tumor-as
135	43	56.6	33	8	ADKS50888	Adks50888 Human C35	208	39	51.3	582	8	ABM81541	Abm81541 Tumor-as
136	43	56.6	33	8	ADKS50800	Adks50800 Human C35	209	39	51.3	582	8	ABM81541	Abm81541 Tumor-as
137	43	56.6	33	8	ADKS50812	Adks50812 Human C35	210	39	51.3	582	8	ABM81541	Abm81541 Tumor-as
138	43	56.6	33	8	ADKS50816	Adks50816 Human C35	211	39	51.3	582	8	ABM81541	Abm81541 Tumor-as
139	43	56.6	33	8	ADKS50824	Adks50824 Human C35	212	39	51.3	582	8	ABM81541	Abm81541 Tumor-as
140	43	56.6	34	8	ADKS50840	Adks50840 Human C35	213	39	51.3	582	8	ABM81541	Abm81541 Tumor-as
141	43	56.6	37	8	ADKS50848	Adks50848 Human C35	214	39	51.3	582	8	ABM81541	Abm81541 Tumor-as
142	43	56.6	38	8	ADKS50884	Adks50884 Human C35	215	39	51.3	582	8	ABM81541	Abm81541 Tumor-as
143	43	56.6	38	8	ADKS50855	Adks50855 Human C35	216	39	51.3	582	8	ABM81541	Abm81541 Tumor-as
144	43	56.6	39	8	ADKS50880	Adks50880 Human C35	217	39	51.3	582	8	ABM81541	Abm81541 Tumor-as
145	43	56.6	39	8	ADKS50828	Adks50828 Human C35	218	39	51.3	582	8	ABM81541	Abm81541 Tumor-as
146	43	56.6	42	8	ADKS50859	Adks50859 Human C35	219	39	51.3	582	8	ABM81541	Abm81541 Tumor-as
147	43	56.6	45	8	ADKS50852	Adks50852 Human C35	220	39	51.3	582	8	ABM81541	Abm81541 Tumor-as
148	43	56.6	49	8	ADKS50863	Adks50863 Human C35	221	39	51.3	582	8	ABM81541	Abm81541 Tumor-as
149	43	56.6	50	8	ADKS50856	Adks50856 Human C35	222	39	51.3	582	8	ABM81541	Abm81541 Tumor-as
150	43	56.6	54	8	ADKS50860	Adks50860 Human C35	223	39	51.3	582	8	ABM81541	Abm81541 Tumor-as
151	43	56.6	54	8	ADKS50867	Adks50867 Human C35	224	39	51.3	582	8	ABM81541	Abm81541 Tumor-as
152	43	56.6	61	8	ADKS50871	Adks50871 Human C35	225	39	51.3	582	8	ABM81541	Abm81541 Tumor-as
153	43	56.6	61	8	ADKS50864	Adks50864 Human C35	226	39	51.3	582	8	ABM81541	Abm81541 Tumor-as
154	43	56.6	66	8	ADKS50868	Adks50868 Human C35	227	39	51.3	582	8	ABM81541	Abm81541 Tumor-as
155	43	56.6	67	8	ADKS50885	Adks50885 Human C35	228	39	51.3	582	8	ABM81541	Abm81541 Tumor-as
156	43	56.6	73	8	ADKS50872	Adks50872 Human C35	229	39	51.3	582	8	ABM81541	Abm81541 Tumor-as
157	43	56.6	79	8	ADKS50876	Adks50876 Human C35	230	39	51.3	582	8	ABM81541	Abm81541 Tumor-as
158	43	56.6	386	5	Aae22205	Aae22205 His6-NLS-	231	39	51.3	582	8	ABM81541	Abm81541 Tumor-as
159	43	56.6	456	7	ABO71756	AbO71756 Pseudomon	232	39	51.3	582	8	ABM81541	Abm81541 Tumor-as
160	42	55.3	20	4	AAE12479	Aae12479 Membrane	233	38	50.0	582	8	ABM81541	Abm81541 Tumor-as
161	42	55.3	78	4	AAW85532	Aaw85532 Human imm	234	38	50.0	582	8	ABM81541	Abm81541 Tumor-as
162	41	53.9	19	3	AAW67956	Aaw67956 Karpesi s	235	38	50.0	582	8	ABM81541	Abm81541 Tumor-as
163	41	53.9	22	3	AAW67958	Aaw67958 Karpesi s	236	38	50.0	582	8	ABM81541	Abm81541 Tumor-as
164	41	53.9	30	2	AAW75651	Aaw75651 Human pla	237	38	50.0	582	8	ABM81541	Abm81541 Tumor-as
165	41	53.9	58	4	AAU53275	Aau53275 Propionib	238	38	50.0	582	8	ABM81541	Abm81541 Tumor-as
166	41	53.9	58	6	ABM49794	Abm49794 Propionib	239	38	50.0	582	8	ABM81541	Abm81541 Tumor-as
167	41	53.9	126	3	AAU53275	Aau53275 Propionib	240	38	50.0	582	8	ABM81541	Abm81541 Tumor-as
168	41	53.9	421	7	ABO70346	AbO70346 Pseudomon	241	38	50.0	582	8	ABM81541	Abm81541 Tumor-as
169	41	53.9	528	2	AAW98295	Aaw98295 H. pylori	242	38	50.0	582	8	ABM81541	Abm81541 Tumor-as
170	41	53.9	528	5	ABO7348	AbO7348 H. pylori	243	38	50.0	582	8	ABM81541	Abm81541 Tumor-as
171	41	53.9	528	6	ABU30791	Abu30791 Protein e	244	38	50.0	582	8	ABM81541	Abm81541 Tumor-as

245	38	50.0	30	8	ADJ45705	Adj45705 Signal pe	318	37	48.7	270	3	AA76332	AA76332 Fragment
246	38	50.0	30	8	ADN43307	Adn43307 Amino aci	319	37	48.7	270	7	ADE11909	Ade11909 Human sec
247	38	50.0	30	8	ADN43274	Adn43274 Amino aci	320	37	48.7	279	2	AY25159	Ay25159 Maize RIP
248	38	50.0	33	8	ADJ45736	Adj45736 Signal pe	321	37	48.7	279	2	AY25160	Ay25160 Maize RIP
249	38	50.0	33	8	ADJ45739	Adj45739 Signal pe	322	37	48.7	281	7	ADG73091	Adg73091 Pseudomon
250	38	50.0	33	8	ADJ45740	Adj45740 Signal pe	323	37	48.7	281	7	ADL12146	Adl12146 Pseudomon
251	38	50.0	71	6	AAU36204	Aau36204 Pseudomon	324	37	48.7	286	4	AA881117	Aa881117 Mycobacte
252	38	50.0	71	6	ABU411550	Abu411550 Protein e	325	37	48.7	287	2	AY25161	Ay25161 Maize RIP
253	38	50.0	71	6	ABU38275	Abu38275 Protein e	326	37	48.7	294	4	ABG27907	Abg27907 Novel hum
254	38	50.0	71	6	ABJ18782	Abj18782 Pseudomon	327	37	48.7	344	6	ABU33804	Abu33804 Protein e
255	38	50.0	71	7	ADC87855	Adc87855 Ribosomal	328	37	48.7	354	2	AA60696	Aa60696 Fragment
256	38	50.0	77	4	AAO12035	Aao12035 Human pol	329	37	48.7	357	7	ABE64996	Abe64996 Human pro
257	38	50.0	81	4	AAW91613	Aaw91613 Human lmm	330	37	48.7	387	2	AY25163	Ay25163 Maize RIP
258	38	50.0	82	6	ADA22840	Ada22840 Triticum	331	37	48.7	421	7	ABO77732	Abo77732 Pseudomon
259	38	50.0	124	6	AAW56732	Aaw56732 Nucellus	332	37	48.7	469	6	ADA54968	Ada54968 Human pro
260	38	50.0	133	8	ABO58584	Abo58584 Human gen	333	37	48.7	479	4	AB92879	Ab92879 Human pro
261	38	50.0	135	7	ABO68686	Abo68686 Pseudomon	334	37	48.7	479	7	ADF76665	Adf76665 Novel hum
262	38	50.0	159	4	AAW25282	Aaw25282 Human pro	335	37	48.7	483	6	ADA89355	Ada89355 DME-10 am
263	38	50.0	232	7	ADC96464	Adc96464 E. faeciu	336	37	48.7	485	7	AAO30822	Aao30822 Human cel
264	38	50.0	253	6	ABM70448	Abm70448 Phototrab	337	37	48.7	499	2	AAW55660	Aaw55660 H. pylori
265	38	50.0	256	3	AAAG31265	Aag31265 Arabidops	338	37	48.7	501	7	ADE28781	Ade28781 Human NOV
266	38	50.0	260	3	AAAG31264	Aag31264 Arabidops	339	37	48.7	501	7	ADE28783	Ade28783 Human NOV
267	38	50.0	287	3	AAAG31263	Aag31263 Arabidops	340	37	48.7	501	7	ADE28779	Ade28779 Human NOV
268	38	50.0	296	8	ADU05170	Adu05170 M. catarr	341	37	48.7	501	8	ADM93524	Adm93524 Human NOV
269	38	50.0	305	6	ABU35394	Abu35394 Protein e	342	37	48.7	501	8	ADM93528	Adm93528 Human NOV
270	38	50.0	309	3	AAAB18898	Aab18898 A maize c	343	37	48.7	501	8	ADM93526	Adm93526 Human NOV
271	38	50.0	330	6	ADA48034	Ada48034 Rice prot	344	37	48.7	505	7	ADE28775	Ade28775 Human NOV
272	38	50.0	357	7	ADG30772	Adg30772 Leptospir	345	37	48.7	505	7	ADE28753	Ade28753 Human NOV
273	38	50.0	370	5	AB807785	Ab807785 E coli mu	346	37	48.7	505	7	ADE28757	Ade28757 Human NOV
274	38	50.0	373	5	AB807786	Ab807786 V cholera	347	37	48.7	505	7	ADE28739	Ade28739 Human NOV
275	38	50.0	378	5	AB807784	Ab807784 E coli he	348	37	48.7	505	7	ADE28743	Ade28743 Human NOV
276	38	50.0	378	6	ABU40405	Abu40405 Protein e	349	37	48.7	505	7	ADE28747	Ade28747 Human NOV
277	38	50.0	418	5	AB877095	Ab877095 Bacillus	350	37	48.7	505	7	ADE28749	Ade28749 Human NOV
278	38	50.0	425	7	ADC12732	Adc12732 Human GPC	351	37	48.7	505	7	ADE28741	Ade28741 Human NOV
279	38	50.0	467	7	ADC96548	Adc96548 E. faeciu	352	37	48.7	505	7	ADE28751	Ade28751 Human NOV
280	38	50.0	515	2	AA14139	Aa14139 Zea mays	353	37	48.7	505	7	ADE28745	Ade28745 Human NOV
281	38	50.0	515	3	AA14139	Aa14139 Zea mays	354	37	48.7	505	7	ADE28755	Ade28755 Human NOV
282	38	50.0	516	2	AAW67617	Aaw67617 P. chryseo	355	37	48.7	505	7	ADE28759	Ade28759 Human NOV
283	38	50.0	548	6	ABU36117	Abu36117 Protein e	356	37	48.7	505	8	ADM93486	Adm93486 Human NOV
284	38	50.0	608	4	ABG211354	Abg211354 Novel hum	357	37	48.7	505	8	ADM93488	Adm93488 Human NOV
285	38	50.0	699	7	ADM26342	Adm26342 Hyperther	358	37	48.7	505	8	ADM93500	Adm93500 Human NOV
286	38	50.0	1114	7	ABO76219	Abo76219 Pseudomon	359	37	48.7	505	8	ADM93498	Adm93498 Human NOV
287	38	50.0	1393	4	AAW78991	Aaw78991 Human pro	360	37	48.7	505	8	ADM93494	Adm93494 Human NOV
288	38	50.0	1677	4	ABG10016	Abg10016 Novel hum	361	37	48.7	505	8	ADM93496	Adm93496 Human NOV
289	38	50.0	1696	7	ADJ69016	Adj69016 Human hea	362	37	48.7	505	8	ADM93502	Adm93502 Human NOV
290	38	50.0	1700	4	AAW39324	Aaw39324 Human pol	363	37	48.7	505	8	ADM93520	Adm93520 Human NOV
291	38	50.0	1720	4	AAW78886	Aaw78886 Human pro	364	37	48.7	505	8	ADM93492	Adm93492 Human NOV
292	38	50.0	1721	4	AAW79870	Aaw79870 Human pro	365	37	48.7	505	8	ADM93490	Adm93490 Human NOV
293	38	50.0	2802	5	ADI16953	Adi16953 Chicken N	366	37	48.7	505	8	ADM93504	Adm93504 Human NOV
294	38	50.0	368	5	ABB48238	Abb48238 Listeria	367	37	48.7	505	8	ADM93484	Adm93484 Human NOV
295	37.5	49.3	368	5	ABB48238	Abb48238 Listeria	368	37	48.7	524	4	AAW39649	Aaw39649 Human pol
296	37	48.7	21	6	AAE35586	Aae35586 TA compri	369	37	48.7	524	7	ADE28737	Ade28737 Human NOV
297	37	48.7	30	2	AAW79386	Aaw79386 DNA encod	370	37	48.7	524	8	ADM93482	Adm93482 Human NOV
298	37	48.7	38	3	AAW76336	Aaw76336 Fragment	371	37	48.7	528	2	AAW55732	Aaw55732 H. pylori
299	37	48.7	38	3	AAW76335	Aaw76335 Fragment	372	37	48.7	528	2	AAW17167	Aaw17167 H. pylori
300	37	48.7	38	3	AAW76335	Aaw76335 Fragment	373	37	48.7	528	5	AB807349	Ab807349 H. pylori
301	37	48.7	62	4	AAO08575	Aao08575 Human pol	374	37	48.7	528	5	AB807347	Ab807347 H. pylori
302	37	48.7	71	4	AAU86516	Aau86516 Novel hum	375	37	48.7	532	6	ABU43564	Abu43564 Protein e
303	37	48.7	71	7	ADB60250	Adb60250 Connectiv	376	37	48.7	543	4	ABB64903	Abb64903 Drosophil
304	37	48.7	80	6	AAE35946	Aae35946 Lolium pe	377	37	48.7	547	4	AAW41435	Aaw41435 Human pol
305	37	48.7	80	6	AAE35937	Aae35937 Lolium pe	378	37	48.7	550	2	AAW98598	Aaw98598 H. pylori
306	37	48.7	82	6	ADA22834	Ada22834 Triticum	379	37	48.7	554	6	ADA54932	Ada54932 Human pro
307	37	48.7	82	6	ADA22849	Ada22849 Triticum	380	37	48.7	561	5	AAW79414	Aaw79414 CADHP-3
308	37	48.7	91	2	AAW55408	Aaw55408 H. pylori	381	37	48.7	561	6	ABR69650	AbR69650 Human CGD
309	37	48.7	97	5	ABP01038	Abp01038 Human ORF	382	37	48.7	562	2	AAW17886	Aaw17886 Phototrab
310	37	48.7	104	4	AAO06543	Aao06543 Human pol	383	37	48.7	562	2	AAW56545	Aaw56545 Toxin Tea
311	37	48.7	109	4	AAU51514	Aau51514 Propionib	384	37	48.7	583	4	AAW75450	Aaw75450 Human aci
312	37	48.7	109	4	AAU95031	Aau95031 Human pro	385	37	48.7	614	6	ABU97386	Abu97386 Amino aci
313	37	48.7	109	6	ABM48033	Abm48033 Propionib	386	37	48.7	649	3	AAAB12140	AaB12140 Hydrophob
314	37	48.7	172	2	AAW25167	Aaw25167 Maize XRI	387	37	48.7	653	6	ABU08083	Abu08083 Human LCC
315	37	48.7	186	6	ABU41298	Abu41298 Protein e	388	37	48.7	669	3	AAW70539	Aaw70539 Human fac
316	37	48.7	191	7	ADF07763	Adf07763 Bacterial	389	37	48.7	729	6	ABU08082	Abu08082 Human LCC
317	37	48.7	248	7	ADC35096	Adc35096 Human bre	390	37	48.7	741	4	AAE05188	Aae05188 Human dru

391	37	48.7	760	5	AAE22157	Aae22157 Human TRN	464	36	47.4	304	6	ABU2089	Novel hum
392	37	48.7	760	7	ADe11758	Human sec	465	36	47.4	304	6	ABJ72269	Human PRO
393	37	48.7	768	5	ABP41137	Human ova	466	36	47.4	304	6	ABJ72397	Human PRO
394	37	48.7	769	6	ADA00877	Rat ESDN.	467	36	47.4	304	6	ABO34292	Human sec
395	37	48.7	769	6	ADA00875	Mouse ESD	468	36	47.4	304	7	ABJ72099	Human mem
396	37	48.7	775	6	ADA00873	Human ESD	469	36	47.4	304	7	ABJ72099	Human mem
397	37	48.7	775	8	ADL61256	Human pro	470	36	47.4	304	7	ADB83574	Novel hum
398	37	48.7	775	8	ABM82701	Human dia	471	36	47.4	304	7	ADB80680	Novel hum
399	37	48.7	794	7	ADP77181	KALPA. 2/	472	36	47.4	304	7	ADB73221	Novel hum
400	37	48.7	810	4	ABG04698	Novel hum	473	36	47.4	304	7	ADB78303	Novel hum
401	37	48.7	828	7	ADe08079	Novel pro	474	36	47.4	304	7	ADB84951	Human PRO
402	37	48.7	1116	7	ADP70293	C. neofor	475	36	47.4	304	7	ADB78057	Novel hum
403	37	48.7	1189	2	AAW17884	Photorhab	476	36	47.4	304	7	ADB87123	Human PRO
404	37	48.7	1189	2	AAW56543	Aaw17884 Toxin Tca	477	36	47.4	304	7	ADB84705	Human PRO
405	37	48.7	1430	4	ABBS58602	Aaw56543 Toxin Tca	478	36	47.4	304	7	ADB83820	Novel hum
406	36.5	48.0	27	2	AAW56393	C-MYCEM	479	36	47.4	304	7	ADB72975	Novel hum
407	36.5	48.0	27	3	AAV55814	C-MYCEM	480	36	47.4	304	7	ADC36813	Human PRO
408	36.5	48.0	86	4	AAO11777	Human pol	481	36	47.4	304	7	ADC21803	Human PRO
409	36	47.4	33	4	AAO12557	Human pol	482	36	47.4	304	7	ADC49834	Novel hum
410	36	47.4	55	5	ABP42101	Human ova	483	36	47.4	304	7	ADC49033	Novel hum
411	36	47.4	57	4	AAW99600	Human exc	484	36	47.4	304	7	ADC49550	Novel hum
412	36	47.4	57	4	AAW43665	Human bla	485	36	47.4	304	7	ADC47411	Novel hum
413	36	47.4	57	8	ADF71624	Human bla	486	36	47.4	304	7	ADC47156	Novel hum
414	36	47.4	62	4	AAU17847	Novel hum	487	36	47.4	304	7	ADC78031	Novel hum
415	36	47.4	62	4	AAU18579	Human lun	488	36	47.4	304	7	ADD06266	Novel hum
416	36	47.4	62	7	ADB33203	Human nov	489	36	47.4	304	7	ADD10655	Human sec
417	36	47.4	62	7	ADG41227	Human res	490	36	47.4	304	7	ADC77785	Novel hum
418	36	47.4	62	7	ADP197001	Human res	491	36	47.4	304	7	ADC77785	Novel hum
419	36	47.4	104	3	AAW53549	Human col	492	36	47.4	304	7	ADD11615	Human sec
420	36	47.4	106	4	AAW74000	Human col	493	36	47.4	304	7	ADD50748	Novel hum
421	36	47.4	127	3	ABO64682	Klebsiell	494	36	47.4	304	7	ADD50994	Novel hum
422	36	47.4	137	3	AAW05116	Arabidops	495	36	47.4	304	7	ADD37408	Human sec
423	36	47.4	145	3	AAW07106	Arabidops	496	36	47.4	304	7	ADD50475	Human PRO
424	36	47.4	145	6	ABU07043	Maize SSI	497	36	47.4	304	7	ADD50229	Human PRO
425	36	47.4	157	4	ABG06106	Novel hum	498	36	47.4	304	7	ADD51240	Novel hum
426	36	47.4	170	6	ABP77932	N. gonorr	499	36	47.4	304	8	ADF59644	Human pol
427	36	47.4	170	6	ABP80045	N. gonorr	500	36	47.4	304	8	ADP78753	G-coupled
428	36	47.4	188	7	ADL33319	Human tra	501	36	47.4	304	8	ADC48787	Novel hum
429	36	47.4	195	5	ABP29636	Streptoco	502	36	47.4	304	8	ADe20958	Novel hum
430	36	47.4	195	6	ABU46888	Protein e	503	36	47.4	304	8	ADe05802	Human PRO
431	36	47.4	206	5	AAO20967	Human 197	504	36	47.4	304	8	ADe05556	Human PRO
432	36	47.4	206	8	ADL42027	Plant tra	505	36	47.4	304	8	ADP75777	Novel hum
433	36	47.4	206	8	ADO02577	Thalecres	506	36	47.4	304	8	ADD85009	Novel hum
434	36	47.4	206	8	ADO62216	Transcrip	507	36	47.4	304	8	ADD86835	Novel hum
435	36	47.4	208	8	ADL98262	Human ser	508	36	47.4	304	8	ADe20712	Novel hum
436	36	47.4	210	4	AAU44679	Propionib	509	36	47.4	304	8	ADe39009	Novel hum
437	36	47.4	210	6	ABW41198	Human ser	510	36	47.4	304	8	ADe05556	Human PRO
438	36	47.4	210	8	ADL98263	Human ser	511	36	47.4	304	8	ADP73541	Human PRO
439	36	47.4	233	8	ADL42025	Plant tra	512	36	47.4	304	8	ADD78381	Novel hum
440	36	47.4	233	8	ADO02575	Thalecres	513	36	47.4	304	8	ADe1616	Human sec
441	36	47.4	233	8	ADO62214	Transcrip	514	36	47.4	304	8	ADe21204	Novel hum
442	36	47.4	250	4	ABG06105	Novel hum	515	36	47.4	304	8	ADD77319	Novel hum
443	36	47.4	260	6	ABU37544	Protein e	516	36	47.4	304	8	ADe20466	Novel hum
444	36	47.4	264	7	ADC08293	Rice prot	517	36	47.4	304	8	ADD75531	Human PRO
445	36	47.4	267	6	ABR48492	Human KGP	518	36	47.4	304	8	ADD74047	Human PRO
446	36	47.4	267	7	ADC08046	Rice prot	519	36	47.4	304	8	ADD74293	Human PRO
447	36	47.4	274	6	ABP78298	N. gonorr	520	36	47.4	304	8	ADD76023	Novel hum
448	36	47.4	275	4	AAE03560	Human dif	521	36	47.4	304	8	ADD78057	Novel hum
449	36	47.4	280	7	ADP07990	Novel pro	522	36	47.4	304	8	ADD75277	Human PRO
450	36	47.4	284	7	ADP58996	Human pol	523	36	47.4	304	8	ADD76821	Novel hum
451	36	47.4	300	6	ABM68711	Photorhab	524	36	47.4	304	8	ADD86589	Novel hum
452	36	47.4	304	3	AAV93650	Amino aci	525	36	47.4	304	8	ADD77565	Novel hum
453	36	47.4	304	4	AAU10060	Insulin-1	526	36	47.4	304	8	ADD77811	Novel hum
454	36	47.4	304	4	ABG06104	Novel hum	527	36	47.4	304	8	ADD85269	Novel hum
455	36	47.4	304	4	ABG06107	Novel hum	528	36	47.4	304	8	ADD73801	Human PRO
456	36	47.4	304	5	AAU83633	Human PRO	529	36	47.4	304	8	ADD74539	Human PRO
457	36	47.4	304	5	ABW84999	Human PRO	530	36	47.4	304	8	ADD77067	Novel hum
458	36	47.4	304	5	AAO20966	Human 304	531	36	47.4	304	8	ADD85761	Novel hum
459	36	47.4	304	5	AAU83226	Novel sec	532	36	47.4	304	8	ADe05310	Human PRO
460	36	47.4	304	5	ABW95605	Human arg	533	36	47.4	304	8	ADD74785	Human PRO
461	36	47.4	304	5	AAE17599	Human ext	534	36	47.4	304	8	ADG05597	Novel hum
462	36	47.4	304	6	ABU80780	Human PRO	535	36	47.4	304	8	ADG27151	Human PRO
463	36	47.4	304	6	ABO33746	Novel hum	536	36	47.4	304	8	ADG11214	Novel hum

537	36	47.4	304	8	ADG11993	Adg11993 Novel hum	610	36	47.4	1001	7	ADB64129	Adb64129 Human pro
538	36	47.4	304	8	ADF94550	Adf94550 Novel hum	611	36	47.4	1055	3	AA119591	Aa119591 Human CAS
539	36	47.4	304	8	ADG06646	Adg06646 Human pro	612	36	47.4	1055	6	ABU07847	Abu07847 Human eph
540	36	47.4	304	8	ADH38990	Adh38990 Novel hum	613	36	47.4	1055	8	ADL61215	Adl61215 Human pro
541	36	47.4	304	8	ADH43799	Adh43799 Human pro	614	36	47.4	1055	8	ADQ89148	Adq89148 Human pro
542	36	47.4	304	8	ADG34080	Adg34080 Novel hum	615	36	47.4	1055	8	ADR44890	Adr44890 Polypepti
543	36	47.4	304	8	ADH33550	Adh33550 Human pro	616	36	47.4	1205	8	ADQ66705	Adq66705 Novel hum
544	36	47.4	304	8	ADH69644	Adh69644 Human pro	617	36	47.4	1310	8	ADN22378	Adn22378 Bacterial
545	36	47.4	304	8	ADH79309	Adh79309 NOV3 prot	618	36	47.4	1722	7	ABM78943	Abm78943 Breast ca
546	36	47.4	304	8	ADI29805	Adi29805 Novel hum	619	36	47.4	1722	8	ADN03808	Adn03808 Antipsori
547	36	47.4	304	8	ADM27202	Adm27202 Novel hum	620	36	47.4	1722	8	ADP23342	Adp23342 PRO polyp
548	36	47.4	304	8	ADR83144	Adr83144 Human pro	621	36	47.4	1909	4	AA11166	Aa11166 Amino aci
549	36	47.4	304	8	ADK66560	Adk66560 Human pro	622	36	47.4	1909	7	ADF75185	Adf75185 Thale cre
550	36	47.4	304	8	ADO55988	Ado55988 Human NOV	623	36	47.4	1923	4	ABG21342	Abg21342 Novel hum
551	36	47.4	304	8	ADSI1108	Adsi1108 Human the	624	36	47.4	4219	8	ADN00372	Adn00372 Novel hum
552	36	47.4	321	8	ADJ76231	Adj76231 Marker ge	625	35.5	46.7	20	4	AAE12495	Aae12495 Membrane
553	36	47.4	321	8	ADJ76200	Adj76200 Marker ge	626	35.5	46.7	22	8	ADJ36110	Adj36110 Self-coal
554	36	47.4	331	7	ADM26706	Adm26706 Hyperther	627	35.5	46.7	23	8	ADJ36116	Adj36116 Self-coal
555	36	47.4	352	6	ABU79148	Abu79148 Chemokine	628	35.5	46.7	309	4	ABG29367	Abg29367 Novel hum
556	36	47.4	352	6	ABU79080	Abu79080 Human che	629	35.5	46.7	830	6	ABU24643	Abu24643 Protein e
557	36	47.4	352	7	ADF43317	Adf43317 Superanti	630	35	46.1	8	2	AAV44161	Aav44161 Truncated
558	36	47.4	352	7	ADF43424	Adf43424 Superanti	631	35	46.1	8	4	AAU00648	Aau00648 Human mem
559	36	47.4	355	8	ADN74359	Adn74359 Thale cre	632	35	46.1	8	7	ABE88791	Abe88791 Membrane
560	36	47.4	358	8	ADS44821	Ads44821 Bacterial	633	35	46.1	8	8	ABG75442	Abg75442 Membrane
561	36	47.4	361	4	ABG07433	Abg07433 Novel hum	634	35	46.1	10	2	AAV44167	Aav44167 Truncated
562	36	47.4	373	8	ADS32637	Ads32637 Sequence	635	35	46.1	10	4	AAU00639	Aau00639 Human mem
563	36	47.4	373	8	ADR73768	Adr73768 Neurospor	636	35	46.1	10	4	AAU00650	Aau00650 Human mem
564	36	47.4	374	6	ABU70702	Abu70702 Human adi	637	35	46.1	10	7	ABE88782	Abe88782 Membrane
565	36	47.4	391	3	AG20311	Ag20311 Arabidops	638	35	46.1	10	7	ABE88793	Abe88793 Membrane
566	36	47.4	391	3	AG61061	Ag61061 Arabidops	639	35	46.1	10	8	ABG75433	Abg75433 Membrane
567	36	47.4	395	8	ABM84893	Abm84893 Human dia	640	35	46.1	11	8	ABG75444	Abg75444 Membrane
568	36	47.4	397	3	AG61060	Ag61060 Arabidops	641	35	46.1	11	8	ADR82253	Adr82253 Membrane
569	36	47.4	398	8	ABM84892	Abm84892 Human dia	642	35	46.1	12	4	AAU00641	Aau00641 Human mem
570	36	47.4	411	6	ABU40835	Abu40835 Protein e	643	35	46.1	12	7	ABE88784	Abe88784 Membrane
571	36	47.4	415	7	ADF04701	Adf04701 Bacterial	644	35	46.1	12	8	ABG75435	Abg75435 Membrane
572	36	47.4	423	7	ABO78467	Abu78467 Pseudomon	645	35	46.1	40	4	AAO13861	Aao13861 Human pol
573	36	47.4	423	8	ADM43142	Adm43142 Human SHT	646	35	46.1	52	4	AAO13861	Aao13861 Human pol
574	36	47.4	424	8	ABM84891	Abm84891 Human dia	647	35	46.1	52	4	AAO13861	Aao13861 Human pol
575	36	47.4	445	8	ADS28516	Ads28516 Bacterial	648	35	46.1	58	4	ABO03961	Abu03961 Human mus
576	36	47.4	477	7	ADS08682	Ads08682 Novel pro	649	35	46.1	58	6	ABU13255	Abu13255 Novel hum
577	36	47.4	477	7	ADS08682	Ads08682 Novel hum	650	35	46.1	58	8	ADJ29281	Adj29281 Human mus
578	36	47.4	478	5	ABE08137	Abu08137 Human S-h	651	35	46.1	59	4	AAU52162	Aau52162 Propionib
579	36	47.4	478	5	ABU09383	Abu09383 Human S-h	652	35	46.1	59	6	ABM48681	Abm48681 Propionib
580	36	47.4	478	5	ABU080480	Abu080480 Ovarian c	653	35	46.1	79	4	ABU12119	Abu12119 Human mem
581	36	47.4	478	7	ABE62796	Abu62796 Human pro	654	35	46.1	87	4	AAO01625	Aao01625 Human pol
582	36	47.4	478	7	ABE61593	Abu61593 Human pro	655	35	46.1	88	5	ABP29239	Abp29239 Streptoco
583	36	47.4	478	7	ABE62792	Abu62792 Human pro	656	35	46.1	91	2	AAU00200	Aau00200 Enterococ
584	36	47.4	478	7	ABE56238	Abu56238 Human pro	657	35	46.1	91	2	AAU00202	Aau00202 Enterococ
585	36	47.4	478	7	ABE61997	Abu61997 Human pro	658	35	46.1	91	5	ABP43421	Abp43421 E faecal
586	36	47.4	478	8	ADJ78509	Adj78509 Human SHT	659	35	46.1	91	5	ABP43419	Abp43419 E faecal
587	36	47.4	478	8	ADM43151	Adm43151 Human SHT	660	35	46.1	91	6	ABU88447	Abu88447 E. faecal
588	36	47.4	478	8	ADL82833	Adl82833 Human pro	661	35	46.1	91	6	ABU88449	Abu88449 E. faecal
589	36	47.4	478	8	ADO05712	Ado05712 Human S-h	662	35	46.1	91	6	ABU13698	Abu13698 Enterococ
590	36	47.4	487	4	ABE60230	Abu60230 Drosophil	663	35	46.1	91	6	ABU13700	Abu13700 Enterococ
591	36	47.4	490	8	ADN20833	Adn20833 Bacterial	664	35	46.1	97	4	ABG03232	Abg03232 Novel hum
592	36	47.4	498	8	ADS41573	Ads41573 Bacterial	665	35	46.1	105	3	ABM42898	Abm42898 Human ORF
593	36	47.4	503	6	ABU33393	Abu33393 Protein e	666	35	46.1	106	2	AAW20103	Aaw20103 H. pylori
594	36	47.4	504	4	ABE52616	Abu52616 Escherich	667	35	46.1	109	5	ABE98891	Abe98891 Human tyr
595	36	47.4	504	4	ABE52616	Abu52616 Escherich	668	35	46.1	112	4	ABG95050	Abg95050 Human pro
596	36	47.4	578	4	ABO69942	Abu69942 Pseudomon	669	35	46.1	113	4	ABG15251	Abg15251 Novel hum
597	36	47.4	592	2	AAW48660	Aaw48660 Human pla	670	35	46.1	114	4	ABG04600	Abg04600 Novel hum
598	36	47.4	592	3	AAW95724	Aaw95724 Cosmid cH	671	35	46.1	114	4	ABG04991	Abg04991 Novel hum
599	36	47.4	695	5	ABU43960	Abu43960 Protein e	672	35	46.1	116	4	ABG15696	Abg15696 Novel hum
600	36	47.4	695	5	ABU43960	Abu43960 Protein e	673	35	46.1	117	8	ADJ48373	Adj48373 Maize oil
601	36	47.4	695	8	ADS29356	Ads29356 Bacterial	674	35	46.1	120	5	ABP42369	Abp42369 Human ova
602	36	47.4	714	4	ABM59870	Abm59870 Drosophil	675	35	46.1	124	2	AAW21899	Aaw21899 Partial 4
603	36	47.4	735	5	ABP35588	Abp35588 Fungal ZB	676	35	46.1	129	7	ADH87679	Adh87679 Enterococ
604	36	47.4	823	4	AAU69746	Aau69746 Human xer	677	35	46.1	130	3	AG27490	Ag27490 Arabidops
605	36	47.4	940	8	ADO03118	Ado03118 Human XPC	678	35	46.1	131	5	ABE49117	Abe49117 Listeria
606	36	47.4	969	8	ADO28627	Ado28627 Human EPH	679	35	46.1	131	6	ABU32896	Abu32896 Protein e
607	36	47.4	977	5	ABU10549	Abu10549 Human NOV	680	35	46.1	132	3	AG24841	Ag24841 Arabidops
608	36	47.4	987	3	AA119590	Aa119590 Human CAS	681	35	46.1	132	3	AG36415	Ag36415 Arabidops
609	36	47.4	987	7	ADN39875	Adn39875 Cancer/an	682	35	46.1	138	4	AAU22155	Aau22155 Human Car

683	35	46.1	138	7	ADB46123	Human car
684	35	46.1	138	8	ADJ07541	Human car
685	35	46.1	139	3	AGG36414	Arabidops
686	35	46.1	139	3	AGG36414	Arabidops
687	35	46.1	139	3	AGG36414	Arabidops
688	35	46.1	139	3	AGG36414	Arabidops
689	35	46.1	139	3	AGG36414	Arabidops
690	35	46.1	139	3	AGG36414	Arabidops
691	35	46.1	139	3	AGG36414	Arabidops
692	35	46.1	139	3	AGG36414	Arabidops
693	35	46.1	139	3	AGG36414	Arabidops
694	35	46.1	139	3	AGG36414	Arabidops
695	35	46.1	139	3	AGG36414	Arabidops
696	35	46.1	139	3	AGG36414	Arabidops
697	35	46.1	139	3	AGG36414	Arabidops
698	35	46.1	139	3	AGG36414	Arabidops
699	35	46.1	139	3	AGG36414	Arabidops
700	35	46.1	139	3	AGG36414	Arabidops
701	35	46.1	139	3	AGG36414	Arabidops
702	35	46.1	139	3	AGG36414	Arabidops
703	35	46.1	139	3	AGG36414	Arabidops
704	35	46.1	139	3	AGG36414	Arabidops
705	35	46.1	139	3	AGG36414	Arabidops
706	35	46.1	139	3	AGG36414	Arabidops
707	35	46.1	139	3	AGG36414	Arabidops
708	35	46.1	139	3	AGG36414	Arabidops
709	35	46.1	139	3	AGG36414	Arabidops
710	35	46.1	139	3	AGG36414	Arabidops
711	35	46.1	139	3	AGG36414	Arabidops
712	35	46.1	139	3	AGG36414	Arabidops
713	35	46.1	139	3	AGG36414	Arabidops
714	35	46.1	139	3	AGG36414	Arabidops
715	35	46.1	139	3	AGG36414	Arabidops
716	35	46.1	139	3	AGG36414	Arabidops
717	35	46.1	139	3	AGG36414	Arabidops
718	35	46.1	139	3	AGG36414	Arabidops
719	35	46.1	139	3	AGG36414	Arabidops
720	35	46.1	139	3	AGG36414	Arabidops
721	35	46.1	139	3	AGG36414	Arabidops
722	35	46.1	139	3	AGG36414	Arabidops
723	35	46.1	139	3	AGG36414	Arabidops
724	35	46.1	139	3	AGG36414	Arabidops
725	35	46.1	139	3	AGG36414	Arabidops
726	35	46.1	139	3	AGG36414	Arabidops
727	35	46.1	139	3	AGG36414	Arabidops
728	35	46.1	139	3	AGG36414	Arabidops
729	35	46.1	139	3	AGG36414	Arabidops
730	35	46.1	139	3	AGG36414	Arabidops
731	35	46.1	139	3	AGG36414	Arabidops
732	35	46.1	139	3	AGG36414	Arabidops
733	35	46.1	139	3	AGG36414	Arabidops
734	35	46.1	139	3	AGG36414	Arabidops
735	35	46.1	139	3	AGG36414	Arabidops
736	35	46.1	139	3	AGG36414	Arabidops
737	35	46.1	139	3	AGG36414	Arabidops
738	35	46.1	139	3	AGG36414	Arabidops
739	35	46.1	139	3	AGG36414	Arabidops
740	35	46.1	139	3	AGG36414	Arabidops
741	35	46.1	139	3	AGG36414	Arabidops
742	35	46.1	139	3	AGG36414	Arabidops
743	35	46.1	139	3	AGG36414	Arabidops
744	35	46.1	139	3	AGG36414	Arabidops
745	35	46.1	139	3	AGG36414	Arabidops
746	35	46.1	139	3	AGG36414	Arabidops
747	35	46.1	139	3	AGG36414	Arabidops
748	35	46.1	139	3	AGG36414	Arabidops
749	35	46.1	139	3	AGG36414	Arabidops
750	35	46.1	139	3	AGG36414	Arabidops

## ALIGNMENTS

RESULT 1	
AAU00654	
ID AAU00654 standard; peptide; 17 AA.	
XX	
AC AAU00654;	
XX	
DT 07-SEP-2001 (first entry)	
XX	
DE Human membrane translocating peptide (MTLP) #23.	
XX	
KW Membrane translocating peptide; MTLP; human; intracellular gene delivery; epithelial cell layer; gastrointestinal tract; circulatory system.	
XX	
OS Homo sapiens.	
XX	
WO200127154-A2.	
XX	
19-APR-2001.	
XX	
27-SEP-2000; 2000WO-IB001491.	
XX	
27-SEP-1999; 99US-0156246P.	
PR	
(OMAH/) O'MAHONY D J.	
PA (LAMB/) LAMBKIN I J.	
XX	
PI O'mahony DJ, Lambkin IJ;	
XX	
WPI; 2001-300212/31.	
DR N-PSDB; AAS00648.	
XX	
Compositions for enhancing uptake of e.g. drugs or DNA across a cell membrane, comprise membrane translocating peptides having specific amino acid sequences or a derivative, fragment, motif, analog or peptidomimetic of the peptides.	
XX	
Claim 2; Page 11; 42pp; English.	
XX	
The sequence represents a human membrane translocated peptide (MTLP). MTLPs and their related fragments, motifs, derivatives and analogues are used for enhancing uptake of a pharmaceutically active agent into a cell, into or out of an intracellular compartment and across a cell layer (for example, an epithelial cell layer lining the gastrointestinal tract), either directly or from a pharmaceutically active agent loaded particle, into the circulatory system of an animal. This method is useful for intracellular gene delivery, as a rapid screening method for the identification of MTLPs which retain the functional activity of a full-length MTLP, as a cell-based screen for assaying the functional activity of a MTLP and characterising the properties of a MTLP, for diagnosis of a pathological disorder (by administration of a MTLP-active agent complex or MTLP-active particle complex comprising a diagnostic agent) and for preventing or treating a pathological disorder	
XX	
SQ Sequence 17 AA;	
Query Match 100.0%; Score 76; DB 4; Length 17;	
Best Local Similarity 100.0%; Pred. No. 1.3e-05; Mismatches 0; Gaps 0;	
Matches 17; Conservative 0; Indels 0;	
QY 1 KKKRKAALVLLPVL 17	
DB 1 KKKRKAALVLLPVL 17	
RESULT 2	
ADB88797	
ID ADB88797 standard; peptide; 17 AA.	
XX	
AC ADB88797;	
XX	
DT 04-DEC-2003 (first entry)	
XX	
DE Membrane translocating peptide #23.	

XX Peyer's patch cell; non-Peyer's patch cell; translocation factor;  
KW upregulated protein; antigen; vaccine delivery; M cell;  
KW membrane translocating peptide.  
XX Unidentified.  
OS  
XX WO2003004646-A2.  
PN  
XX 16-JAN-2003.  
PD  
XX  
XX  
XX 04-APR-2002; 2002WO-IB003866.  
PF  
XX 04-APR-2001; 2001US-0281387P.  
PR  
XX 02-JUL-2001; 2001US-0302591P.  
PR  
XX (OMAH/) O'MAHONY D J.  
PA  
XX O'mahony DJ, Byrne D, Brayden D;  
PI WPI; 2003-229409/22.  
PI  
XX  
XX Increasing the levels of a protein in a Peyer's patch cell, useful for  
PT targeted vaccine or drug delivery, comprises delivering to the Peyer's  
PT patch cell a transcription factor or an activator of a transcription  
PT factor.  
XX  
XX Example 6; Page 51; 147pp; English.  
PS  
XX The invention relates to a novel method for increasing the levels of a  
CC protein in a Peyer's patch cell. The method comprises delivering to the  
CC cell a nucleic acid coding for a protein, the level of which or its mRNA  
CC is greater than in a non-Peyer's patch cell. The preferred protein of the  
CC invention is a transcription factor or a protein that activates a  
CC transcription factor selected from Jun-B, c-Jun related TF, Jun-D, STAT 3  
CC -signal transducer and activator of transcription 3, NkappaBgr; Tf p105  
CC subunit, S-myc proto-oncogene, myc related, Nm23-W2, nucleoside  
CC diphosphate kinase B, metastasis reducing protein, and C-est-I proto-  
CC oncogene, and p54. The preferred upregulated protein of the invention is  
CC selected from clusterin, T-cell surface glycoprotein CD5 precursor, HSP  
CC 84, and C22 dependent phospholipase A2 precursor and the mRNA is for a  
CC protein selected from the group. The method is useful for increasing or  
CC decreasing the level of a protein in a Peyer's patch cell, particularly  
CC in increasing antigen or vaccine delivery to M cells. The method may also  
CC be used to enhance transport of a drug through the gastrointestinal tract  
CC (GIT). This sequence represents a membrane translocating peptide of the  
CC invention.  
XX  
XX Sequence 17 AA;  
SQ  
Query Match 100.0%; Score 76; DB 7; Length 17;  
Best Local Similarity 100.0%; Pred. No. 1.3e-05;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Oy 1 KKKKKAAAALLPVL 17  
Db 1 KKKKKAAAALLPVL 17  
RESULT 3  
ABG75448  
ID ABG75448 standard; peptide; 17 AA.  
XX  
XX ABG75448;  
AC  
XX  
XX 15-APR-2004 (first entry)  
DT  
XX Membrane translocating peptide #23.  
DE  
XX Membrane translocating peptide; transmembrane transport;  
KW enzymatic degradation; gastrointestinal tract.  
XX  
XX Unidentified.  
OS

XX Key Location/Qualifiers  
FH modified\_site 2  
FT /note= "modified by epsilon-dansyl"  
FT modified\_site 17  
FT /note= "C-terminal amide"  
XX  
XX WO2003089458-A2.  
PN  
XX 30-OCT-2003.  
PD  
XX  
XX 21-APR-2003; 2003WO-US012543.  
PF  
XX 19-APR-2002; 2002US-00126845.  
PR  
XX (SARL-) SARLAN LTD.  
PA  
XX O'mahony DJ, Lambkin I, Houghten R, Pinilla C;  
PI WPI; 2004-034528/03.  
PI  
XX  
XX Composition, useful for treating a pathological disease in an animal,  
PT comprises a translocating peptide, consisting of a transport peptide, an  
PT extended peptide comprising the transport peptide or a transport-active  
PT fragment.  
XX  
XX Claim 1; Page 25; Opp; English.  
PS  
XX The present invention relates to a composition which comprises a  
CC translocating peptide, consisting of a transport peptide, an extended  
CC peptide comprising the transport peptide or a transport-active fragment  
CC of at least 4 amino acids of the transport peptide. The composition is  
CC useful in treating a pathological disorder in an animal, as it protects  
CC an orally delivered active agent from enzymatic degradation in the  
CC gastrointestinal tract (GIT), and promotes absorption across epithelial  
CC cells lining the GIT. The present sequence is a polypeptide used in the  
CC exemplification of the invention  
XX  
XX Sequence 17 AA;  
SQ  
Query Match 100.0%; Score 76; DB 8; Length 17;  
Best Local Similarity 100.0%; Pred. No. 1.3e-05;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Oy 1 KKKKKAAAALLPVL 17  
Db 1 KKKKKAAAALLPVL 17  
RESULT 4  
AAV55835  
ID AAV55835 standard; peptide; 29 AA.  
XX  
XX AAV55835;  
AC  
XX  
XX 06-MAR-2000 (first entry)  
DT  
XX  
XX SV40MEM polypeptide fragment.  
DE  
XX  
XX Cellular protein; nuclear translocation; nuclear localization signal;  
KW immunosuppressant; immune response; viral infection; immune disorder;  
KW rheumatoid arthritis; multiple sclerosis; juvenile-onset diabetes; NLS;  
KW systemic lupus erythematosus; SLE; asthma; sepsis; tumor growth; SV40;  
KW fibroblast growth factor; SV40 large T antigen.  
XX  
XX Synthetic.  
OS  
XX Simian virus 40.  
OS  
XX WO9957138-A1.  
PN  
XX 11-NOV-1999.  
PD  
XX  
XX 26-APR-1999; 99WO-US008984.  
PF

XX 04-MAY-1998; 98US-00072429.  
 XX (BRIM ) BRISTOL-MYERS SQUIBB CO.  
 PA Nadler SG;  
 PI  
 XX WPI; 2000-062141/05.  
 XX Novel composition comprising a nuclear localization signal and an  
 PT inhibitor of nuclear translocation, useful for preventing transplanted  
 PT organ or tissue rejection.  
 XX  
 XX Claim 11; Page 42; 68pp; English.  
 XX  
 CC The invention provides a composition comprising an isolated polypeptide  
 CC (I) (polypeptide inhibitor of a cellular protein nuclear translocation)  
 CC which comprises a signal sequence peptide capable of delivering (I)  
 CC through a cytoplasmic membrane into the cell; and a nuclear localization  
 CC signal sequence (NLS) present along with an immunosuppressant. (I) or its  
 CC derivatives provide useful tools for introducing an exogenous polypeptide  
 CC comprising an NLS into an intact cell to inhibit nuclear translocation of  
 CC a cellular protein, for studying the role of nuclear translocation in the  
 CC regulation of cellular processes. See AAY55812 for detailed uses of (I)  
 CC and compositions containing (I). The present sequence represents a  
 CC polypeptide fragment comprising the hydrophobic region of the signal  
 CC peptide from fibroblast growth factor and the SV40 large T antigen NLS  
 XX Sequence 29 AA;  
 SQ

Query Match 70.4%; Score 53.5; DB 3; Length 29;  
 Best Local Similarity 77.8%; Pred. No. 0.12;  
 Matches 14; Conservative 0; Mismatches 3; Indels 1; Gaps 1;  
 QY 1 KKKRKAATAAVLLP-VLLA 17  
 Db 2 KKKRKVAVALLPVALLA 19  
 ||||| || |||||  
 ||||| || |||||

RESULT 5  
 ABU69594  
 ID ABU69594 standard; peptide; 29 AA.  
 XX  
 AC ABU69594;  
 XX  
 XX 05-JUN-2003 (first entry)  
 XX Human NF-kappaB inhibitory peptide.  
 XX  
 XX Human; nuclear factor-kappaB; NF-kappaB; immune disorder; cancer;  
 KW inflammatory disorder; apoptosis; hepatic disorder; Hodgkin's lymphoma;  
 KW haematopoietic tumour; hyper-IGM syndrome; viral infection; asthma;  
 KW hypohidrotic ectodermal dysplasia; human immunodeficiency virus; HIV;  
 KW X-linked anhidrotic ectodermal dysplasia; al incontinentia pigmenti;  
 KW influenza; rheumatoid arthritis; inflammatory bowel disease; colitis;  
 KW atherosclerosis; cachexia; euthyroid sick syndrome; stroke; EAE;  
 KW experimental allergic encephalomyelitis; autoimmune disorder; wound;  
 KW hyper immune activity; acute phase response; hypercongenital condition;  
 KW birth defect; necrotic lesion; organ transplant rejection; pancreas;  
 KW signal transduction; hyperproliferative disorder; diabetes mellitus;  
 KW vitamin B12 malabsorption; neurological disorder; Huntington's chorea;  
 KW Turner's syndrome; bacterial infection; cardiovascular disorder;  
 KW infertility; psoriasis; haemolytic anaemia; antiinflammatory; anti-HIV;  
 KW cytostatic; hepatotropic; virucide; anti rheumatic; antiarthritic;  
 KW antidiabetic; immunomodulator; antidiabetic; anti allergic;  
 KW neuroprotective; immunosuppressive; vulnerary; antibacterial;  
 KW antifertility; antianaemic; antipsoriatic; cerebroprotective; cardiac;  
 KW antiarteriosclerotic.  
 XX  
 XX Homo sapiens.  
 OS  
 XX WO200286076-A2.  
 PN  
 XX

PD 31-OCT-2002.  
 XX 19-APR-2002; 2002WO-US012636.  
 PF  
 XX 19-APR-2001; 2001US-0284962P.  
 PR 26-APR-2001; 2001US-0286645P.  
 PR 09-JAN-2002; 2002US-0346986P.  
 XX  
 XX (BRIM ) BRISTOL-MYERS SQUIBB CO.  
 PA Carman J, Feder J, Nadler S;  
 XX WPI; 2003-093119/08.  
 XX  
 XX Novel NF-kappaB-associated polypeptides and polynucleotides useful for  
 PT diagnosing, treating and preventing cancer, hepatic disorders, aberrant  
 PT apoptosis, viral infections, autoimmune disorders, asthma and stroke.  
 XX  
 PS Disclosure; Fig 1; 608pp; English.  
 XX  
 CC The present invention relates to the isolation of human nuclear factor-  
 CC kappaB (NF-kappaB) associated polypeptides and polynucleotides. The NF-  
 CC kappaB associated polypeptide and polynucleotide sequences are useful for  
 CC preventing, treating or ameliorating various disorders including immune  
 CC disorders, inflammatory disorders, cancers, disorders relating to  
 CC aberrant apoptosis, hepatic disorders, Hodgkin's lymphomas, disorders  
 CC haematopoietic tumours, hyper-IGM syndromes, hypohidrotic ectodermal  
 CC dysplasia, X-linked anhidrotic ectodermal dysplasia, immunodeficiency, al  
 CC incontinentia pigmenti, viral infections (e.g. those caused by human  
 CC immunodeficiency virus (HIV), human T-cell lymphotropic virus (HTLV),  
 CC hepatitis B, hepatitis C, Epstein Barr virus (EBV), influenza),  
 CC rheumatoid arthritis, inflammatory bowel disease, colitis, asthma,  
 CC atherosclerosis, cachexia, euthyroid sick syndrome, stroke, experimental  
 CC allergic encephalomyelitis (EAE), autoimmune disorders, disorders related  
 CC to hyper immune activity, disorders related to aberrant acute phase  
 CC responses, hypercongenital conditions, birth defects, necrotic lesions,  
 CC wounds, organ transplant rejection, disorders related to aberrant signal  
 CC transduction, hyperproliferative disorders, diseases of the pancreas  
 CC (e.g. diabetes mellitus, vitamin B12 malabsorption), neurological  
 CC disorders (e.g. Huntington's chorea), Turner's syndrome, bacterial  
 CC infections, cardiovascular disorders, infertility, psoriasis and  
 CC haemolytic anaemia. The present sequence represents a human NF-kappaB  
 CC inhibitory peptide used to identify the NF-kappaB associated  
 CC polynucleotides of the invention  
 XX Sequence 29 AA;  
 SQ

Query Match 70.4%; Score 53.5; DB 6; Length 29;  
 Best Local Similarity 77.8%; Pred. No. 0.12;  
 Matches 14; Conservative 0; Mismatches 3; Indels 1; Gaps 1;  
 QY 1 KKKRKAATAAVLLP-VLLA 17  
 Db 2 KKKRKVAVALLPVALLA 19  
 ||||| || |||||  
 ||||| || |||||

RESULT 6  
 AAWS6391  
 ID AAWS6391 standard; peptide; 30 AA.  
 XX  
 AC AAWS6391;  
 XX  
 XX 05-AUG-1998 (first entry)  
 XX SV40MEM polypeptide used to inhibit kappa-Ig light chain expression.  
 XX  
 KW SV40MEM polypeptide; signal peptide; fibroblast growth factor;  
 KW SV40 large antigen; nuclear localisation signal; NLS;  
 KW immunosuppressive activity; inhibition; nuclear translocation inhibitor;  
 KW kappa immunoglobulin light chain expression; S. typhosa LPS;  
 KW nuclear translocation; treatment; immune disorder; autoimmune disease;  
 KW hypersensitivity; sepsis; prevention; septic shock; antiviral agent;  
 KW tumour growth suppressor.



```

XX OS Synthetic.
XX OS Simian virus 40.
XX PN WO9811907-A1.
XX PD 26-MAR-1998.
XX PF 15-SEP-1997; 97WO-US016217.
XX PR 20-SEP-1996; 96US-0026978P.
XX PR 12-SEP-1997; 97US-00928958.
XX PA (BRIM ) BRISTOL-MYERS SQUIBB CO.
XX PI Nadler SG, Cleaveland JS, Blake J, Haffar OK;
XX PI WPI; 1998-217028/19.
XX DR Nuclear translocation inhibitor polypeptides - comprising signal sequence
XX PT for delivery through the cytoplasmic membrane and at least 2 nuclear
XX PT localisation sequences.
XX PS Claim 15; Page 28; 69pp; English.
XX CC The present sequence represents the SV40MEM polypeptide, which comprises
XX CC the hydrophobic region of the signal peptide from fibroblast growth
XX CC factor and the SV40 large antigen nuclear localisation signal (NLS). The
XX CC peptide has immunosuppressive activity, and causes an approximately 75-
XX CC 80% inhibition of kappa immunoglobulin (Ig) light chain expression in
XX CC response to S. typhosa LPS. L- and D-forms of the SV40MEM peptide are
XX CC equally effective. The SV40MEM polypeptide exemplifies the nuclear
XX CC translocation inhibitor polypeptides comprise a signal sequence peptide
XX CC capable of delivering the polypeptide through the cytoplasmic membrane
XX CC into a cell, and at least 2 nuclear localisation sequences (NLSs). The
XX CC polypeptides can be used to inhibit nuclear translocation of a cellular
XX CC protein. In addition, since the nuclear translocation of certain cellular
XX CC peptides is required for the host organism to mount an immune response,
XX CC the polypeptide inhibitors are useful as immunosuppression agents. The
XX CC polypeptides can therefore be used for the treatment of immune disorders
XX CC including autoimmune diseases. The polypeptides can also be used for
XX CC treating physical symptoms manifested by responses to allergens which can
XX CC initiate a state of hypersensitivity, for the treatment of sepsis and in
XX CC the prevention of septic shock, antiviral agents, tumour growth
XX CC suppressors, and for transcriptionally modulating the expression of
XX CC cellular genes
XX SQ Sequence 30 AA;

Query Match 70.4%; Score 53.5; DB 2; Length 30;
Best Local Similarity 77.8%; Pred. No. 0.12;
Matches 14; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

Oy 1 KKKRKAARVLLP-VLLA 17
Db 2 KKKRKAARVALLPAVLLA 19

RESULT 7
AAY55812
ID AAY55812 standard; peptide; 30 AA.
XX AC AAY55812;
XX DT 06-MAR-2000 (first entry)
XX DE SV40MEM polypeptide fragment.
XX KW Cellular protein; nuclear translocation; nuclear localization signal;
XX KW immunosuppressant; immune response; viral infection; immune disorder;
XX KW rheumatoid arthritis; multiple sclerosis; juvenile-onset diabetes; NLS;
XX KW systemic lupus erythematosus; SLE; asthma; sepsis; tumor growth; SV40;
XX KW fibroblast growth factor; SV40 large T antigen.

```

```

XX OS Synthetic.
XX OS Simian virus 40.
XX PN WO9957138-A1.
XX PD 11-NOV-1999.
XX PF 26-APR-1999; 99WO-US008984.
XX PR 04-MAY-1998; 98US-00072429.
XX PA (BRIM ) BRISTOL-MYERS SQUIBB CO.
XX PI Nadler SG;
XX PI WPI; 2000-062141/05.
XX DR Novel composition comprising a nuclear localization signal and an
XX PT inhibitor of nuclear translocation, useful for preventing transplanted
XX PT organ or tissue rejection.
XX PS Example 1; Page 26; 68pp; English.

```

The invention provides a composition comprising an isolated polypeptide (I) (polypeptide inhibitor of a cellular protein nuclear translocation) which comprises a signal sequence peptide capable of delivering (I) through a cytoplasmic membrane into the cell; and a nuclear localization signal sequence (NLS) present along with an immunosuppressant. (I) or its derivatives provide useful tools for introducing an exogenous polypeptide comprising an NLS into an intact cell to inhibit nuclear translocation of a cellular protein, for studying the role of nuclear translocation in the regulation of cellular processes. (I) along with other immunosuppressants are useful as immunosuppression compositions, used for preventing or suppressing an immune response i.e. rejection of transplanted organ or tissue in a subject, preferably mammals. Administration of antiviral amounts of the composition into a subject is useful for treating or preventing a viral infection. The composition also is useful in the treatment of a wide variety of immune disorders such as rheumatoid arthritis, multiple sclerosis, juvenile-onset diabetes, systemic lupus erythematosus (SLE), autoimmune uveoretinitis, autoimmune oophoritis, Sjogren's syndrome, granulomatous orchitis, autoimmune vasculitis, Crohn's disease, sarcoidosis, rheumatic carditis, ankylosing spondylitis, Grave's disease, and autoimmune thrombocytopenic purpura as well as to prevent the rejection of transplanted tissue or organs. They are also useful for treating physical symptoms manifested by response to allergens which can initiate a state of hypersensitivity, or which can provoke a hypersensitivity reaction in a subject already sensitized with all allergens. Such physical symptoms include asthma, joint swelling and urticaria. Additionally, due to the superior immunosuppressive properties, the compositions of the present invention are useful in the treatment of sepsis and in the prevention of septic shock. They are also used as antiviral agents and also can be used to suppress tumor growth. (I) and the immunosuppressant work synergistically and provide better immune suppression than either treatment alone. The composition exhibits superior immunosuppressive characteristics compared to using an immunosuppressant or polypeptide alone. The present sequence represents SV40MEM polypeptide fragment comprising the hydrophobic region of the signal peptide from fibroblast growth factor and the SV40 large T antigen NLS

Sequence 30 AA;

Query Match 70.4%; Score 53.5; DB 3; Length 30;  
Best Local Similarity 77.8%; Pred. No. 0.12;  
Matches 14; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

Oy 1 KKKRKAARVLLP-VLLA 17  
Db 2 KKKRKAARVALLPAVLLA 19

RESULT 8

AAU00633  
ID AAU00633 standard; peptide; 16 AA.  
XX  
AC AAU00633;  
XX  
DT 07-SEP-2001 (first entry)  
XX  
XX Human membrane translocating peptide (MTLP) #2.  
DE  
XX Membrane translocating peptide; MTLP; human; intracellular gene delivery;  
KW epithelial cell layer; gastrointestinal tract; circulatory system.  
XX  
XX Homo sapiens.  
XX  
XX WO200127154-A2.  
PN  
XX 19-APR-2001.  
PD  
XX 27-SEP-2000; 2000WO-1B001491.  
XX  
XX 27-SEP-1999; 99US-0156246P.  
XX  
XX (OMAH/) O'MAHONY D J.  
PA (LAMB/) LAMBKIN I J.  
XX  
XX O'mahony DJ, Lambkin IJ;  
PI  
XX  
XX WPI; 2001-300212/31.  
DR N-PSDB; AAS00627.  
XX  
XX Compositions for enhancing uptake of e.g. drugs or DNA across a cell  
PT membrane, comprise membrane translocating peptides having specific amino  
PT acid sequences or a derivative, fragment, motif, analog or peptidomimetic  
PT of the peptides.  
XX  
XX Claim 2; Page 11; 42pp; English.  
XX  
XX The sequence represents a human membrane translocated peptide (MTLP).  
CC MTLPs and their related fragments, motifs, derivatives and analogues are  
CC used for enhancing uptake of a pharmaceutically active agent into a cell,  
CC into or out of an intracellular compartment and across a cell layer (for  
CC example, an epithelial cell layer lining the gastrointestinal tract),  
CC either directly or from a pharmaceutically active agent loaded particle,  
CC into the circulatory system of an animal. This method is useful for  
CC intracellular gene delivery, as a rapid screening method for the  
CC identification of MTLPs which retain the functional activity of a full-  
CC length MTLP, as a cell-based screen for assaying the functional activity  
CC of a MTLP and characterising the properties of a MTLP, for diagnosis of a  
CC pathological disorder (by administration of a MTLP-active agent complex  
CC or MTLP-active particle complex comprising a diagnostic agent) and for  
CC preventing or treating a pathological disorder  
XX  
XX Sequence 16 AA;  
SQ

Query Match 69.7%; Score 53; DB 4; Length 16;  
Best Local Similarity 85.7%; Pred. No. 0.076;  
Matches 12; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 4 RKAAAVLLPVLLA 17  
DB 1 KKKAAVLLPVLLA 14

RESULT 9  
ABP70215  
ID ABP70215 standard; peptide; 16 AA.  
XX  
XX ABP70215;  
AC  
XX  
XX 07-APR-2003 (first entry)  
DT  
XX Amino acid sequence of membrane translocating peptide Zelan094.  
DE  
XX

KW Lipid-nucleic acid complex; polycation; targeting factor; gene therapy;  
KW cancer; infection; immune deficiency; gene defect; genetic disease;  
KW membrane translocating peptide.  
XX  
XX Unidentified.  
XX  
XX WO200288318-A2.  
PN  
XX 07-NOV-2002.  
PD  
XX 30-APR-2002; 2002WO-US013609.  
XX  
XX 30-APR-2001; 2001US-0287786P.  
XX  
XX (TARG-) TARGETED GENETICS CORP.  
PA (EMER-) EMERALD GENE SYSTEMS LTD.  
XX  
XX Harvie P, Paul R, Cudmore S, O'mahony DJ;  
PI  
XX WPI; 2003-183837/18.  
DR  
XX Lipid-nucleic acid complex useful for delivering a nucleic acid to a  
XX cell, comprises compacted nucleic acid, polycation, targeting factor and  
XX lipid, and does not comprise protamine or its salt.  
XX  
XX Disclosure; Page 10; 259pp; English.  
XX  
XX The specification describes a lipid-nucleic acid complex, comprising a  
XX compacted nucleic acid, a polycation, a targeting factor and a lipid, but  
XX not a protamine. The targeting factor increases cellular bioavailability  
XX of the nucleic acid without interaction with a specific outer cell  
XX surface membrane receptor. The mean diameter of the complex is greater  
XX than 100 nm and less than 400 nm. The lipid-nucleic acid complex is  
XX useful for delivering a nucleic acid to a cell in vivo, e.g. for gene  
XX therapy. It reduces levels of inflammatory cytokines such as tumour  
XX necrosis factor-alpha. The complex is useful for manufacturing a  
XX medicament for treating or diagnosing a variety of diseases, conditions  
XX or syndromes such as cancer, bacterial, viral or parasitic infections,  
XX immune deficiencies, gene defects, and gene deficiencies (e.g. inherited  
XX genetic diseases). The present sequence represents a membrane  
XX translocating peptide, which is used as the targeting factor in lipid-  
XX nucleic acid complexes of the invention  
XX  
XX Sequence 16 AA;  
SQ

Query Match 69.7%; Score 53; DB 6; Length 16;  
Best Local Similarity 85.7%; Pred. No. 0.076;  
Matches 12; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 4 RKAAAVLLPVLLA 17  
DB 1 KKKAAVLLPVLLA 14

RESULT 10  
ADB88776  
ID ADB88776 standard; peptide; 16 AA.  
XX  
XX ADB88776;  
AC  
XX  
XX 04-DEC-2003 (first entry)  
DT  
XX  
XX Membrane translocating peptide #2.  
DE  
XX Peyer's patch cell; non-Peyer's patch cell; transcription factor;  
KW upregulated protein; antigen; vaccine delivery; M cell;  
KW membrane translocating peptide.  
XX  
XX Unidentified.  
XX  
XX WO2003004646-A2.  
PN  
XX 16-JAN-2003.  
PD

XX 04-APR-2002; 2002WO-IB003866.  
XX  
XX  
PR 04-APR-2001; 2001US-0281387P.  
PR 02-JUL-2001; 2001US-0302591P.  
XX  
XX (OMAH/) O'MAHONY D J.  
XX  
XX O'mahony DJ, Byrne D, Brayden D;  
XX WPI; 2003-229409/22.  
XX  
XX Increasing the levels of a protein in a Peyer's patch cell, useful for  
PT targeted vaccine or drug delivery, comprises delivering to the Peyer's  
PT patch cell a transcription factor or an activator of a transcription  
PT factor.  
XX  
XX Example 6; Page 51; 147pp; English.  
XX  
XX The invention relates to a novel method for increasing the levels of a  
CC protein in a Peyer's patch cell. The method comprises delivering to the  
CC cell a nucleic acid coding for a protein, the level of which or its mRNA  
CC is greater than in a non-Peyer's patch cell. The preferred protein of the  
CC invention is a transcription factor or a protein that activates a  
CC transcription factor selected from Jun-B, c-Jun related TF, Jun-D, STAT 3  
CC -signal transducer and activator of transcription 3, Nfkapabgr; Tf p105  
CC subunit, S-myc proto-oncogene, myc related, Nm23-M2, nucleoside  
CC diphosphate kinase B, metatasis reducing protein, and C-est-I proto-  
CC oncogene, and p54. The preferred upregulated protein of the invention is  
CC selected from clusterin, T-cell surface glycoprotein CD5 precursor, HSP  
CC 84, and Ca2+ dependent phospholipase A2 precursor and the mRNA is for a  
CC protein selected from the group. The method is useful for increasing or  
CC decreasing the level of a protein in a Peyer's patch cell, particularly  
CC in increasing antigen or vaccine delivery to M cells. The method may also  
CC be used to enhance transport of a drug through the gastrointestinal tract  
CC (GIT). This sequence represents a membrane translocating peptide of the  
CC invention.  
XX  
XX Sequence 16 AA;  
SQ  
Query Match 69.7%; Score 53; DB 7; Length 16;  
Best Local Similarity 85.7%; Pred. No. 0.076;  
Matches 12; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
Oy 4 KKAATAVLLPVLLA 17  
Db 1 KKAATAVLLPVLLA 14  
RESULT 11  
ABP70216  
ID ABP70216 standard; peptide; 20 AA.  
XX  
XX AC ABP70216;  
XX  
XX DT 07-APR-2003 (first entry)  
XX  
XX DE Amino acid sequence of membrane translocating peptide ZELAN094R.  
XX  
XX Lipid-nucleic acid complex; polycation; targeting factor; gene therapy;  
KW cancer; infection; immune deficiency; gene defect; genetic disease;  
KW membrane translocating peptide.  
XX  
XX OS Unidentified.  
XX  
XX FH Key Location/Qualifiers  
XX Modified-site 1  
FT /note= "S(galactose), cholesteryl-succinyl, DOP-  
FT succinyl, DSPE-PEG5K-succinyl, DMPE-PEG5K-succinyl, or  
XX DSPE-PEG5K-succinyl attached"  
XX PN WO200288318-A2.  
XX  
XX PD 07-NOV-2002.  
XX  
XX PF 30-APR-2002; 2002WO-US013609.  
XX  
XX PD 07-NOV-2002.  
XX  
XX PF 30-APR-2002; 2002WO-US013609.  
XX  
XX PR 30-APR-2001; 2001US-0287786P.  
XX  
XX (TARG-) TARGETED GENETICS CORP.  
XX (EMER-) EMERALD GENE SYSTEMS LTD.  
XX Harvie P, Paul R, Cudmore S, O'mahony DJ;

PA (TARG-) TARGETED GENETICS CORP.  
PA (EMER-) EMERALD GENE SYSTEMS LTD.  
XX  
XX Harvie P, Paul R, Cudmore S, O'mahony DJ;  
XX  
XX WPI; 2003-183837/18.  
XX  
XX Lipid-nucleic acid complex useful for delivering a nucleic acid to a  
PT cell, comprises compacted nucleic acid, polycation, targeting factor and  
PT lipid, and does not comprise protamine or its salt.  
XX  
XX Disclosure; Page 10; 259pp; English.  
XX  
XX The specification describes a lipid-nucleic acid complex, comprising a  
CC compacted nucleic acid, a polycation, a targeting factor and a lipid, but  
CC not a protamine. The targeting factor increases cellular bioavailability  
CC of the nucleic acid without interaction with a specific outer cell  
CC surface membrane receptor. The mean diameter of the complex is greater  
CC than 100 nm and less than 400 nm. The lipid-nucleic acid complex is  
CC useful for delivering a nucleic acid to a cell in vivo, e.g. for gene  
CC therapy. It reduces levels of inflammatory cytokines such as tumour  
CC necrosis factor-alpha. The complex is useful for manufacturing a  
CC medicament for treating or diagnosing a variety of diseases, conditions  
CC or syndromes such as cancer, bacterial, viral or parasitic infections,  
CC immune deficiencies, gene defects, and gene deficiencies (e.g. inherited  
CC genetic diseases). The present sequence represents a membrane  
CC translocating peptide, which is used as the targeting factor in lipid-  
CC nucleic acid complexes of the invention  
XX  
XX Sequence 20 AA;  
SQ  
Query Match 69.7%; Score 53; DB 6; Length 20;  
Best Local Similarity 85.7%; Pred. No. 0.096;  
Matches 12; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
Oy 4 KKAATAVLLPVLLA 17  
Db 1 KKAATAVLLPVLLA 14  
RESULT 12  
ABP70234  
ID ABP70234 standard; peptide; 15 AA.  
XX  
XX AC ABP70234;  
XX  
XX DT 07-APR-2003 (first entry)  
XX  
XX DE Amino acid sequence of targeting factor Gelan094.  
XX  
XX Lipid-nucleic acid complex; polycation; targeting factor; gene therapy;  
KW cancer; infection; immune deficiency; gene defect; genetic disease.  
XX  
XX OS Unidentified.  
XX  
XX FH Key Location/Qualifiers  
XX Modified-site 1  
FT /note= "S(galactose), cholesteryl-succinyl, DOP-  
FT succinyl, DSPE-PEG5K-succinyl, DMPE-PEG5K-succinyl, or  
XX DSPE-PEG5K-succinyl attached"  
XX PN WO200288318-A2.  
XX  
XX PD 07-NOV-2002.  
XX  
XX PF 30-APR-2002; 2002WO-US013609.  
XX  
XX PD 07-NOV-2002.  
XX  
XX PF 30-APR-2001; 2001US-0287786P.  
XX  
XX (TARG-) TARGETED GENETICS CORP.  
XX (EMER-) EMERALD GENE SYSTEMS LTD.  
XX  
XX Harvie P, Paul R, Cudmore S, O'mahony DJ;

XX WPI; 2003-183837/18.  
XX Lipid-nucleic acid complex useful for delivering a nucleic acid to a  
PT cell, comprises compacted nucleic acid, polycation, targeting factor and  
PT lipid, and does not comprise protamine or its salt.  
XX Disclosure; Page 42; 259pp; English.  
XX The specification describes a lipid-nucleic acid complex, comprising a  
CC compacted nucleic acid, a polycation, a targeting factor and a lipid, but  
CC not a protamine. The targeting factor increases cellular bioavailability  
CC of the nucleic acid. The targeting factor increases cellular bioavailability  
CC of the nucleic acid without interaction with a specific outer cell  
CC surface membrane receptor. The mean diameter of the complex is greater  
CC than 100 nm and less than 400 nm. The lipid-nucleic acid complex is  
CC useful for delivering a nucleic acid to a cell in vivo, e.g. for gene  
CC therapy. It reduces levels of inflammatory cytokines such as tumour  
CC necrosis factor-alpha. The complex is useful for manufacturing a  
CC medicament for treating or diagnosing a variety of diseases, conditions  
CC or syndromes such as cancer, bacterial, viral or parasitic infections,  
CC immune deficiencies, gene defects, and gene deficiencies (e.g. inherited  
CC genetic diseases). The present sequence represents a targeting factor for  
CC lipid-nucleic acid complexes of the invention  
XX  
XX Sequence 15 AA;  
SQ  
Query Match 67.1%; Score 51; DB 6; Length 15;  
Best Local Similarity 92.3%; Pred. No. 0.15;  
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 5 KAAAVLLPVLLA 17  
DB 1 KAAAVLLPVLLA 13  
RESULT 13  
ABP70214  
ID ABP70214 standard; peptide; 15 AA.  
XX  
XX ABP70214;  
AC  
XX 07-APR-2003 (first entry)  
DT  
XX  
XX Amino acid sequence of membrane translocating peptide Elan094.  
DE  
XX  
XX Lipid-nucleic acid complex; polycation; targeting factor; gene therapy;  
KW cancer; infection; immune deficiency; gene defect; genetic disease;  
KW membrane translocating peptide.  
XX  
XX Unidentified.  
OS  
XX  
XX WO200288318-A2.  
PN  
XX  
XX 07-NOV-2002.  
PD  
XX  
XX 30-APR-2002; 2002WO-US013609.  
PF  
XX  
XX 07-NOV-2002.  
PD  
XX  
XX 30-APR-2002; 2002WO-US013609.  
PF  
XX  
XX 30-APR-2001; 2001US-0287786P.  
PR  
XX  
XX (TARG-) TARGETED GENETICS CORP.  
PA  
XX (EMER-) EMERALD GENE SYSTEMS LTD.  
PA  
XX  
XX Harvie P, Paul R, Cudmore S, O'mahony DJ;  
PI  
XX WPI; 2003-183837/18.  
DR  
XX  
XX Lipid-nucleic acid complex useful for delivering a nucleic acid to a  
PT cell, comprises compacted nucleic acid, polycation, targeting factor and  
PT lipid, and does not comprise protamine or its salt.  
XX  
XX Disclosure; Page 10; 259pp; English.  
PS  
XX  
XX The specification describes a lipid-nucleic acid complex, comprising a

CC compacted nucleic acid, a polycation, a targeting factor and a lipid, but  
CC not a protamine. The targeting factor increases cellular bioavailability  
CC of the nucleic acid without interaction with a specific outer cell  
CC surface membrane receptor. The mean diameter of the complex is greater  
CC than 100 nm and less than 400 nm. The lipid-nucleic acid complex is  
CC useful for delivering a nucleic acid to a cell in vivo, e.g. for gene  
CC therapy. It reduces levels of inflammatory cytokines such as tumour  
CC necrosis factor-alpha. The complex is useful for manufacturing a  
CC medicament for treating or diagnosing a variety of diseases, conditions  
CC or syndromes such as cancer, bacterial, viral or parasitic infections,  
CC immune deficiencies, gene defects, and gene deficiencies (e.g. inherited  
CC genetic diseases). The present sequence represents a membrane  
CC translocating peptide, which is used as the targeting factor in lipid-  
CC nucleic acid complexes of the invention  
XX  
XX Sequence 15 AA;  
SQ  
Query Match 67.1%; Score 51; DB 6; Length 15;  
Best Local Similarity 92.3%; Pred. No. 0.15;  
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 5 KAAAVLLPVLLA 17  
DB 1 KAAAVLLPVLLA 13  
RESULT 14  
ABP70233  
ID ABP70233 standard; peptide; 15 AA.  
XX  
XX ABP70233;  
AC  
XX 07-APR-2003 (first entry)  
DT  
XX  
XX Amino acid sequence of targeting factor Elan094G.  
DE  
XX  
XX Lipid-nucleic acid complex; polycation; targeting factor; gene therapy;  
KW cancer; infection; immune deficiency; gene defect; genetic disease.  
KW  
XX  
XX Unidentified.  
OS  
XX  
XX Key Location/Qualifiers  
FH  
XX Modified-site 1  
FT /note= "Des-Pro attached"  
FT Modified-site 15  
FT /note= "galactose attached"  
FT  
XX WO200288318-A2.  
PN  
XX  
XX 07-NOV-2002.  
PD  
XX  
XX 30-APR-2002; 2002WO-US013609.  
PF  
XX  
XX 30-APR-2001; 2001US-0287786P.  
PR  
XX  
XX (TARG-) TARGETED GENETICS CORP.  
PA  
XX (EMER-) EMERALD GENE SYSTEMS LTD.  
PA  
XX  
XX Harvie P, Paul R, Cudmore S, O'mahony DJ;  
PI  
XX WPI; 2003-183837/18.  
DR  
XX  
XX Lipid-nucleic acid complex useful for delivering a nucleic acid to a  
PT cell, comprises compacted nucleic acid, polycation, targeting factor and  
PT lipid, and does not comprise protamine or its salt.  
XX  
XX Disclosure; Page 42; 259pp; English.  
PS  
XX  
XX The specification describes a lipid-nucleic acid complex, comprising a  
CC compacted nucleic acid, a polycation, a targeting factor and a lipid, but  
CC not a protamine. The targeting factor increases cellular bioavailability  
CC of the nucleic acid without interaction with a specific outer cell  
CC surface membrane receptor. The mean diameter of the complex is greater

than 100 nm and less than 400 nm. The lipid-nucleic acid complex is useful for delivering a nucleic acid to a cell in vivo, e.g. for gene therapy. It reduces levels of inflammatory cytokines such as tumour necrosis factor- $\alpha$ . The complex is useful for manufacturing a medicament for treating or diagnosing a variety of diseases, conditions or syndromes such as cancer, bacterial, viral or parasitic infections, immune deficiencies, gene defects, and gene deficiencies (e.g. inherited genetic diseases). The present sequence represents a targeting factor for lipid-nucleic acid complexes of the invention

SQ Sequence 15 AA;

Query Match 67.1%; Score 51; DB 6; Length 15;  
Best Local Similarity 92.3%; Pred. No. 0.15;  
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 KAAAVLLPVLLA 17  
| | | | | | | | | |  
Db 1 KAAAVLLPVLLA 13

RESULT 15  
ABG75426  
ID ABG75426 standard; peptide; 15 AA.

XX AC ABG75426;

XX DT 15-APR-2004 (first entry)

XX DE Membrane translocating peptide #1.

XX KW Membrane translocating peptide; transmembrane transport; enzymatic degradation; gastrointestinal tract.

XX OS Unidentified.

XX FH Key Location/Qualifiers

FT modified\_site 2 /note= "modified by epsilon-dansyl"

FT modified\_site 15 /note= "C-terminal amide"

XX WO2003089458-A2.

XX PD 30-OCT-2003.

XX PF 21-APR-2003; 2003WO-US012543.

XX PR 19-APR-2002; 2002US-00126845.

XX PA (SARL-) SARLAN LTD.

XX PI O'mahony DJ, Lambkin I, Houghten R, Pinilla C;

XX DR WPI; 2004-034528/03.

XX Composition, useful for treating a pathological disease in an animal, comprises a translocating peptide, consisting of a transport peptide, an extended peptide comprising the transport peptide or a transport-active fragment of at least 4 amino acids of the transport peptide. The composition is useful in treating a pathological disorder in an animal, as it protects an orally delivered active agent from enzymatic degradation in the gastrointestinal tract (GIT), and promotes absorption across epithelial cells lining the GIT. The present sequence is a polypeptide used in the exemplification of the invention

XX Claim 1; Page 24; Opp; English.

XX The present invention relates to a composition which comprises a translocating peptide, consisting of a transport peptide, an extended peptide comprising the transport peptide or a transport-active fragment of at least 4 amino acids of the transport peptide. The composition is useful in treating a pathological disorder in an animal, as it protects an orally delivered active agent from enzymatic degradation in the gastrointestinal tract (GIT), and promotes absorption across epithelial cells lining the GIT. The present sequence is a polypeptide used in the exemplification of the invention

SQ Sequence 15 AA;

Query Match 67.1%; Score 51; DB 8; Length 15;  
Best Local Similarity 92.3%; Pred. No. 0.15;  
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 KAAAVLLPVLLA 17  
| | | | | | | | | |  
Db 1 KAAAVLLPVLLA 13

RESULT 16

AAU00655

ID AAU00655 standard; peptide; 19 AA.

XX AC AAU00655;

XX DT 07-SEP-2001 (first entry)

XX DE Human targeting peptide sequence #1.

XX KW Membrane translocating peptide; MTLP; human; intracellular gene delivery; epithelial cell layer; gastrointestinal tract; circulatory system; targeting peptide.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

FT Modified-site 1 /note= "N-terminus has a dansyl group"

FT Modified-site 19 /note= "C-terminal amide"

XX WO200127154-A2.

XX PD 19-APR-2001.

XX PF 27-SEP-2000; 2000WO-IB001491.

XX PR 27-SEP-1999; 99US-0156246P.

XX PA (OMAH)/ O'MAHONY D J.

XX PA (LAMB)/ LAMBKIN I J.

XX PI O'mahony DJ, Lambkin IJ;

XX DR WPI; 2001-300212/31.

XX Compositions for enhancing uptake of e.g. drugs or DNA across a cell membrane, comprise membrane translocating peptides having specific amino acid sequences or a derivative, fragment, motif, analog or peptidomimetic of the peptides.

XX Example 1; Page 22; 42pp; English.

XX The sequence represents a human targeting peptide which is used to identify the presence of a human membrane translocated peptide (MTLP) and quantify the amount, to bind the MTLP to the surface of a particle, or to localise the MTLP in a cell or tissue sample. MTLPs and their related fragments, motifs, derivatives and analogues are used for enhancing uptake of a pharmaceutically active agent into a cell, into or out of an intracellular compartment and across a cell layer (for example, an epithelial cell layer lining the gastrointestinal tract), either directly or from a pharmaceutically active agent loaded particle, into the circulatory system of an animal. This method is useful for intracellular gene delivery, as a rapid screening method for the identification of MTLPs which retain the functional activity of a full-length MTLP, as a cell-based screen for assaying the functional activity of a MTLP and characterising the properties of a MTLP, for diagnosis of a pathological disorder (by administration of a MTLP-active agent complex or MTLP-active particle complex comprising a diagnostic agent) and for preventing or treating a pathological disorder

## SQ Sequence 19 AA;

Query Match 67.1%; Score 51; DB 4; Length 19;  
Best Local Similarity 92.3%; Pred. No. 0.19;  
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 KAAAVLLPVLIA 17  
| | | | | | | | | |  
Db 1 KKAANVLLPVLIA 13

## RESULT 17

AAU00634  
ID AAU00634 standard; peptide; 19 AA.

XX AC AAU00634;

XX DT 07-SEP-2001 (first entry)

XX DE Human membrane translocating peptide (MTLP) #3.

XX KW Membrane translocating peptide; MTLP; human; intracellular gene delivery;  
epithelial cell layer; gastrointestinal tract; circulatory system.

XX OS Homo sapiens.

XX PN WO200127154-A2.

XX PD 19-APR-2001.

XX PF 27-SEP-2000; 2000WO-IB001491.

XX PR 27-SEP-1999; 99US-0156246P.

XX PA (OMAH/) O'MAHONY D J.

XX PA (LAMB/) LAMBKIN I J.

XX PI O'mahony DJ, Lambkin IJ;

XX DR WPI; 2001-300212/31.

XX DR N-PSDB; AAS00628.

XX PT Compositions for enhancing uptake of e.g. drugs or DNA across a cell  
membrane, comprise membrane translocating peptides having specific amino  
acid sequences or a derivative, fragment, motif, analog or peptidomimetic  
of the peptides.

XX PS Claim 2; Page 11; 42pp; English.

CC The sequence represents a human membrane translocated peptide (MTLP).  
CC MTLPs and their related fragments, motifs, derivatives and analogues are  
CC used for enhancing uptake of a pharmaceutically active agent into a cell,  
CC into or out of an intracellular compartment and across a cell layer (for  
CC example, an epithelial cell layer lining the gastrointestinal tract),  
CC either directly or from a pharmaceutically active agent loaded particle,  
CC into the circulatory system of an animal. This method is useful for  
CC intracellular gene delivery, as a rapid screening method for the  
CC identification of MTLPs which retain the functional activity of a full-  
CC length MTLP, as a cell-based screen for assaying the functional activity  
CC of a MTLP and characterising the properties of a MTLP, for diagnosis of a  
CC pathological disorder (by administration of a MTLP-active agent complex  
CC or MTLP-active particle complex comprising a diagnostic agent) and for  
CC preventing or treating a pathological disorder

## SQ Sequence 19 AA;

Query Match 67.1%; Score 51; DB 4; Length 19;  
Best Local Similarity 92.3%; Pred. No. 0.19;  
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 KAAAVLLPVLIA 17  
| | | | | | | | | |  
Db 1 KKAANVLLPVLIA 13

## RESULT 18

ADB88777  
ID ADB88777 standard; peptide; 19 AA.

XX AC ADB88777;

XX DT 04-DEC-2003 (first entry)

XX DE Membrane translocating peptide #3.

XX KW Peyer's patch cell; non-Peyer's patch cell; transcription factor;  
upregulated protein; antigen; vaccine delivery; M cell;  
membrane translocating peptide.

XX OS Unidentified.

XX PN WO2003004646-A2.

XX PD 16-JAN-2003.

XX PF 04-APR-2002; 2002WO-IB003866.

XX PR 04-APR-2001; 2001US-0281387P.

XX PR 02-JUL-2001; 2001US-0302591P.

XX PA (OMAH/) O'MAHONY D J.

XX PI O'mahony DJ, Byrne D, Brayden D;

XX DR WPI; 2003-229409/22.

XX PT Increasing the levels of a protein in a Peyer's patch cell, useful for  
targeted vaccine or drug delivery, comprises delivering to the Peyer's  
patch cell a transcription factor or an activator of a transcription  
factor.

XX PS Example 6; Page 51; 147pp; English.

XX CC The invention relates to a novel method for increasing the levels of a  
protein in a Peyer's patch cell. The method comprises delivering to the  
cell a nucleic acid coding for a protein, the level of which or its mRNA  
is greater than in a non-Peyer's patch cell. The preferred protein of the  
invention is a transcription factor or a protein that activates a  
transcription factor selected from Jun-B, C-jun related TF, Jun-D, STAT 3  
-signal transducer and activator of transcription 3, Nfkapppabgr; Tf p105  
subunit, S-myc proto-oncogene, myc related, Nm23-M2, nucleoside  
diphosphate kinase B, metastasis reducing protein, and C-est-1 proto-  
oncogene, and p54. The preferred upregulated protein of the invention is  
selected from clusterin, T-cell surface glycoprotein CD5 precursor, HSP  
84, and Ca2+ dependent phospholipase A2 precursor and the mRNA is for a  
protein selected from the group. The method is useful for increasing or  
decreasing the level of a protein in a Peyer's patch cell, particularly  
in increasing antigen or vaccine delivery to M cells. The method may also  
be used to enhance transport of a drug through the gastrointestinal tract  
(GIT). This sequence represents a membrane translocating peptide of the  
invention.

## SQ Sequence 19 AA;

Query Match 67.1%; Score 51; DB 7; Length 19;  
Best Local Similarity 92.3%; Pred. No. 0.19;  
Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 5 KAAAVLLPVLIA 17  
| | | | | | | | | |  
Db 1 KKAANVLLPVLIA 13

## RESULT 19

ABG75428  
ID ABG75428 standard; peptide; 19 AA.

```

XX AC ABG75428;
XX DT 15-APR-2004 (first entry)
XX DE Membrane translocating peptide #3.
XX KW Membrane translocating peptide; transmembrane transport;
XX KW enzymatic degradation; gastrointestinal tract.
XX OS Unidentified.
XX FH Key Location/Qualifiers
FT modified_site 2
FT /note= "modified by epsilon-dansyl"
FT modified_site 19
FT /note= "C-terminal amide"
XX WO2003089458-A2.
XX PN 30-OCT-2003.
XX PD
XX PF 21-APR-2003; 2003WO-US012543.
XX PR 19-APR-2002; 2002US-00126845.
XX PA (SARL-) SARLAN LTD.
XX PI O'mahony DJ, Lambkin I, Houghten R, Pinilla C;
XX WPI; 2004-034528/03.
XX PT Composition, useful for treating a pathological disease in an animal,
XX PT comprises a translocating peptide, consisting of a transport peptide, an
XX PT extended peptide comprising the transport peptide or a transport-active
XX PT fragment.
XX PS Claim 1; Page 24; Opp: English.
XX CC The present invention relates to a composition which comprises a
XX CC translocating peptide, consisting of a transport peptide, an extended
XX CC peptide comprising the transport peptide or a transport-active fragment
XX CC of at least 4 amino acids of the transport peptide. The composition is
XX CC useful in treating a pathological disorder in an animal, as it protects
XX CC an orally delivered active agent from enzymatic degradation in the
XX CC gastrointestinal tract (GIT), and promotes absorption across epithelial
XX CC cells lining the GIT. The present sequence is a polypeptide used in the
XX CC exemplification of the invention
XX SQ Sequence 19 AA;
XX Query Match 67.1%; Score 51; DB 8; Length 19;
XX Best Local Similarity 92.3%; Pred. No. 0.19;
XX Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 5 KAAAVLLPVLIA 17
DB 1 KKAARVLLPVLIA 13
XX
RESULT 20
AAU00632
ID AAU00632 standard; peptide; 21 AA.
XX AC AAU00632;
XX DT 07-SEP-2001 (first entry)
XX DE Human membrane translocating peptide (MTLP) #1.
XX KW Membrane translocating peptide; MTLP; human; intracellular gene delivery;
XX KW epithelial cell layer; gastrointestinal tract; circulatory system.
XX
OS Homo sapiens.
XX PN WO200127154-A2.
XX XX
XX PD 19-APR-2001.
XX XX
XX PF 27-SEP-2000; 2000WO-IB001491.
XX PR 27-SEP-1999; 99US-0156246P.
XX PA (OMAH/) O'MAHONY D J.
XX PA (LAMB/) LAMBKIN I J.
XX PI O'mahony DJ, Lambkin IJ;
XX WPI; 2001-300212/31.
XX DR N-PSDB; AAS00626.
XX XX
XX PT Compositions for enhancing uptake of e.g. drugs or DNA across a cell
XX PT membrane, comprise membrane translocating peptides having specific amino
XX PT acid sequences or a derivative, fragment, motif, analog or peptidomimetic
XX PT of the peptides.
XX PS Claim 1; Page 11; 42pp; English.
XX XX
XX CC The sequence represents a human membrane translocated peptide (MTLP).
XX CC MTLPs and their related fragments, motifs, derivatives and analogues are
XX CC used for enhancing uptake of a pharmaceutically active agent into a cell,
XX CC into or out of an intracellular compartment and across a cell layer (for
XX CC example, an epithelial cell layer lining the gastrointestinal tract),
XX CC either directly or from a pharmaceutically active agent loaded particle,
XX CC into the circulatory system of an animal. This method is useful for
XX CC intracellular gene delivery, as a rapid screening method for the
XX CC identification of MTLPs which retain the functional activity of a full-
XX CC length MTLP, as a cell-based screen for assaying the functional activity
XX CC of a MTLP and characterising the properties of a MTLP, for diagnosis of a
XX CC pathological disorder (by administration of a MTLP-active agent complex
XX CC or MTLP-active particle complex comprising a diagnostic agent) and for
XX CC preventing or treating a pathological disorder
XX SQ Sequence 21 AA;
XX Query Match 67.1%; Score 51; DB 4; Length 21;
XX Best Local Similarity 92.3%; Pred. No. 0.22;
XX Matches 12; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 5 KAAAVLLPVLIA 17
DB 1 KKAARVLLPVLIA 13
XX
RESULT 21
ADB88775
ID ADB88775 standard; peptide; 21 AA.
XX AC ADB88775;
XX DT 04-DEC-2003 (first entry)
XX DE Membrane translocating peptide #1.
XX KW Peyer's patch cell; non-Peyer's patch cell; transcription factor;
XX KW upregulated protein; antigen; vaccine delivery; M cell;
XX KW membrane translocating peptide.
XX OS Unidentified.
XX PN WO2003004646-A2.
XX PD 16-JAN-2003.
XX PF 04-APR-2002; 2002WO-IB003866.
XX XX

```





PT Novel C35 polypeptide useful for formulation of immunogenic composition  
 to induce antibodies and cell-mediated immunity against tumor cells.  
 PS Disclosure; Page 225; 626pp; English.

XX The invention relates to a novel isolated polypeptide comprising or  
 CC consisting of two or more C35 peptide epitopes. The polypeptide of the  
 CC invention demonstrates cytostatic activity and may be useful for the  
 CC formulation of an immunogenic composition, such as a vaccine, to induce  
 CC antibodies and cell-mediated immunity against target cells such as tumour  
 CC cells. Furthermore, the polypeptide and its analogues may be useful as  
 CC prognostic markers for carcinoma, such as human breast or bladder  
 CC carcinoma. The current sequence is that of human C35 / Kaposi FGF  
 CC (fibroblast growth factor) signal sequence MST (membrane-translocating  
 CC sequence) cell-penetrating polypeptide of the invention.

XX Sequence 21 AA;

Query Match 65.8%; Score 50; DB 8; Length 21;  
 Best Local Similarity 85.7%; Pred. No. 0.31;  
 Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 RKAAAVLLPVLLA 17  
 | | | | | | | | | |  
 DB 6 RPPAAVLLPVLLA 19

RESULT 24

ADK50920  
 ID ADK50920 standard; peptide; 33 AA.

XX ADK50920;

DT 04-NOV-2004 (first entry)

XX Human C35 / Kaposi FGF signal sequence MST cell-penetrating peptide 132.

XX C35 epitope; cytostatic; vaccine; tumour; breast; bladder carcinoma;  
 KW human; Kaposi FGF signal sequence h region MST; cell-penetrating;  
 KW fibroblast growth factor; membrane-translocating signal.

XX Homo sapiens.  
 OS Unidentified.

XX WO2003104428-A2.

XX 18-DEC-2003.

PF 10-JUN-2003; 2003WO-US018252.

XX 10-JUN-2002; 2002US-0386738P.

PR 11-DEC-2002; 2002US-0432241P.

PR 23-APR-2003; 2003US-0464650P.

XX (VACC-) VACCINEX INC.

PA (UYRP) UNIV ROCHESTER.

XX Zauderer M, Evans EE, Borrello MA;

XX WPI; 2004-062349/06.

XX Novel C35 polypeptide useful for formulation of immunogenic composition  
 to induce antibodies and cell-mediated immunity against tumor cells.

XX Disclosure; Page 225; 626pp; English.

XX The invention relates to a novel isolated polypeptide comprising or  
 CC consisting of two or more C35 peptide epitopes. The polypeptide of the  
 CC invention demonstrates cytostatic activity and may be useful for the  
 CC formulation of an immunogenic composition, such as a vaccine, to induce  
 CC antibodies and cell-mediated immunity against target cells such as tumour  
 CC cells. Furthermore, the polypeptide and its analogues may be useful as  
 CC prognostic markers for carcinoma, such as human breast or bladder

CC carcinoma. The current sequence is that of human C35 / Kaposi FGF  
 CC (fibroblast growth factor) signal sequence MST (membrane-translocating  
 CC sequence) cell-penetrating polypeptide of the invention.

XX Sequence 33 AA;

Query Match 65.8%; Score 50; DB 8; Length 33;  
 Best Local Similarity 85.7%; Pred. No. 0.5;  
 Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4 RKAAAVLLPVLLA 17  
 | | | | | | | | | |  
 DB 18 RPPAAVLLPVLLA 31

RESULT 25

ADP54374  
 ID ADP54374 standard; protein; 433 AA.

XX ADP54374;

XX 18-NOV-2004 (first entry)

XX Human PRO protein sequence SEQ ID NO:350.

XX human; PRO; immune related disease; inflammatory immune response;  
 KW immune response stimulation; anti-allergic; antianaemic; antiarthritic;  
 KW antiasthmatic; antidiabetic; antiinflammatory; antipsoriatic;  
 KW antirheumatic; antithyroid; CNS; dermatological; gastrointestinal;  
 KW haemostatic; hepatotropic; immunostimulant; immunosuppressive; muscular;  
 KW nephrotropic; neuroprotective; osteopathic; respiratory; vasotropic;  
 KW virucide; gene therapy.

XX Homo sapiens.

XX WO2004039956-A2.

XX 13-MAY-2004.

XX 28-OCT-2003; 2003WO-US034381.

XX 29-OCT-2002; 2002US-0422472P.

XX (GETH) GENENTECH INC.

XX Aggarwal S, Clark H, Gurney AL, Schoenfeld J, Williams PM;  
 PI Wood WI, Wu TD;

XX WPI; 2004-376182/35.

XX N-PSDB; ADP54373.

XX New PRO polynucleotides and polypeptides, useful in useful in diagnosing  
 PT and treating an immune related disease, e.g. systemic lupus  
 PT erythematosus, rheumatoid arthritis, diabetes mellitus or asthma and in  
 PT stimulating an immune response.

XX Claim 1; SEQ ID NO 350; 3009pp; English.

XX The present invention describes an isolated PRO nucleic acid (1). Also  
 CC described: (1) a vector comprising (1); (2) a host cell comprising the  
 CC vector of (1); (3) a process for producing a PRO polypeptide; (4) an  
 CC isolated PRO polypeptide; (5) a chimeric molecule comprising the  
 CC polypeptide of (4) fused to a heterologous amino acid sequence; (6) an  
 CC antibody which specifically binds to a polypeptide of (4); (7) a  
 CC composition of matter comprising a polypeptide of (4), an agonist or  
 CC antagonist of the polypeptide or an antibody that binds to the  
 CC polypeptide in combination with a carrier; (8) an article of manufacture  
 CC comprising a container, a label on the container and a composition of  
 CC matter of (7); (9) a method of treating an immune related disease in a  
 CC mammal; (10) a method for determining the presence of a PRO polypeptide  
 CC in a sample suspected of having the polypeptide; (11) a method of  
 CC diagnosing an immune related disease or an inflammatory immune response  
 CC in mammal; (12) a method of identifying a compound that inhibits or



CC into the circulatory system of an animal. This method is useful for  
CC intracellular gene delivery, as a rapid screening method for the  
CC identification of MTLPs which retain the functional activity of a full-  
CC length MTLP, as a cell-based screen for assaying the functional activity  
CC of a MTLP and characterising the properties of a MTLP, for diagnosis of a  
CC pathological disorder (by administration of a MTLP-active agent complex  
CC or MTLP-active particle complex comprising a diagnostic agent) and for  
CC preventing or treating a pathological disorder

XX Sequence 14 AA;

Query Match 61.8%; Score 47; DB 4; Length 14;  
Best Local Similarity 84.6%; Pred. No. 0.64;  
Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 KAAAVLLPVLLA 17  
DB 1 KKCAAVLLPVLLA 13  
|||||

RESULT 28

ADB88780  
ID ADB88780 standard; peptide; 14 AA.

XX AC ADB88780;

DT 04-DEC-2003 (first entry)

DE Membrane translocating peptide #6.

XX Peyer's patch cell; non-Peyer's patch cell; transcription factor;  
KW upregulated protein; antigen; vaccine delivery; M cell;  
KW membrane translocating peptide.

XX Unidentified.

XX WO2003004646-A2.

XX 16-JAN-2003.

XX 04-APR-2002; 2002WO-IB003866.

XX 04-APR-2001; 2001US-0281387P.

PR 02-JUL-2001; 2001US-0302591P.

XX (OMAH/) O'MAHONY D J.

XX O'mahony DJ, Byrne D, Brayden D;

PI WPI; 2003-229409/22.

XX Increasing the levels of a protein in a Peyer's patch cell, useful for  
PT targeted vaccine or drug delivery, comprises delivering to the Peyer's  
PT patch cell a transcription factor or an activator of a transcription  
PT factor.

XX Example 6; Page 51; 147pp; English.

XX The invention relates to a novel method for increasing the levels of a  
CC protein in a Peyer's patch cell. The method comprises delivering to the  
CC cell a nucleic acid coding for a protein, the level of which or its mRNA  
CC is greater than in a non-Peyer's patch cell. The preferred protein of the  
CC invention is a transcription factor or a protein that activates a  
CC transcription factor selected from Jun-B, c-jun related TF, Jun-D, STAT 3  
CC -signal transducer and activator of transcription 3, Nfkapabgr; Tf p105  
CC subunit, S-myc proto-oncogene, myc related, Nm23-M2, nucleoside  
CC diphosphate kinase B, metastasis reducing protein, and C-est-I proto-  
CC oncogene, and p54. The preferred upregulated protein of the invention is  
CC selected from clusterin, T-cell surface glycoprotein CD5 precursor, HSP  
CC 84, and Ca2+ dependent phospholipase A2 precursor and the mRNA is for a  
CC protein selected from the group. The method is useful for increasing or  
CC decreasing the level of a protein in a Peyer's patch cell, particularly  
CC in increasing antigen or vaccine delivery to M cells. The method may also

CC be used to enhance transport of a drug through the gastrointestinal tract  
CC (GIT). This sequence represents a membrane translocating peptide of the  
CC invention.

XX Sequence 14 AA;

Query Match 61.8%; Score 47; DB 7; Length 14;  
Best Local Similarity 84.6%; Pred. No. 0.64;  
Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 5 KAAAVLLPVLLA 17  
DB 1 KKCAAVLLPVLLA 13  
|||||

RESULT 29

ABG75427  
ID ABG75427 standard; peptide; 14 AA.

XX AC ABG75427;

XX 15-APR-2004 (first entry)

XX Membrane translocating peptide #2.

XX Membrane translocating peptide; transmembrane transport;  
KW enzymatic degradation; gastrointestinal tract.

XX Unidentified.

XX Key Location/Qualifiers  
FH modified\_site 14  
FT /note= "C-terminal amide"

XX WO2003089458-A2.

XX 30-OCT-2003.

XX 21-APR-2003; 2003WO-US012543.

XX 19-APR-2002; 2002US-00126845.

XX (SARL-) SARLAN LTD.

XX O'mahony DJ, Lambkin I, Houghten R, Pinilla C;

XX WPI; 2004-034528/03.

XX Composition, useful for treating a pathological disease in an animal,  
PT comprises a translocating peptide, consisting of a transport peptide, an  
PT extended peptide comprising the transport peptide or a transport-active  
PT fragment.

XX Claim 1; Page 24; Opp; English.

XX The present invention relates to a composition which comprises a  
CC translocating peptide, consisting of a transport peptide, an extended  
CC peptide comprising the transport peptide or a transport-active fragment  
CC of at least 4 amino acids of the transport peptide. The composition is  
CC useful in treating a pathological disorder in an animal, as it protects  
CC an orally delivered active agent from enzymatic degradation in the  
CC gastrointestinal tract (GIT), and promotes absorption across epithelial  
CC cells lining the GIT. The present sequence is a polypeptide used in the  
CC exemplification of the invention

XX Sequence 14 AA;

Query Match 61.8%; Score 47; DB 8; Length 14;  
Best Local Similarity 100.0%; Pred. No. 0.64;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 AAAVLLPVLLA 17  
|||||

```

Db          2  AAAVLLPVLLA 12

RESULT 30
ABG75431
ID  ABG75431 standard; peptide; 14 AA.
XX
AC  ABG75431;
XX
DT  15-APR-2004 (first entry)
XX
DE  Membrane translocating peptide #6.
XX
KW  Membrane translocating peptide; transmembrane transport;
KW  enzymatic degradation; gastrointestinal tract.
XX
OS  Unidentified.
XX
FH  Key Location/Qualifiers
FT  modified_site 2 /note= "modified by epsilon-dansyl"
FT  modified_site 14
FT  modified_site 14 /note= "C-terminal amide"
XX
XX  WO2003089458-A2.
XX
XX  30-OCT-2003.
XX
XX  21-APR-2003; 2003WO-US012543.
XX
XX  19-APR-2002; 2002US-00126845.
XX
XX  (SARL-) SARLAN LTD.
XX
XX  O'mahony DJ, Lambkin I, Houghten R, Pinilla C;
XX  WPI; 2004-034528/03.
XX
XX  Composition, useful for treating a pathological disease in an animal,
XX  comprises a translocating peptide, consisting of a transport peptide, an
XX  extended peptide comprising the transport peptide or a transport-active
XX  fragment.
XX
XX  Claim 1; Page 24; 0pp; English.
XX
XX  The present invention relates to a composition which comprises a
XX  translocating peptide, consisting of a transport peptide, an extended
XX  peptide comprising the transport peptide or a transport-active fragment
XX  of at least 4 amino acids of the transport peptide. The composition is
XX  useful in treating a pathological disorder in an animal, as it protects
XX  an orally delivered active agent from enzymatic degradation in the
XX  gastrointestinal tract (GIT), and promotes absorption across epithelial
XX  cells lining the GIT. The present sequence is a polypeptide used in the
XX  exemplification of the invention
XX
SQ  Sequence 14 AA;
    Query Match 61.8%; Score 47; DB 8; Length 14;
    Best Local Similarity 84.6%; Pred. No. 0.64;
    Matches 11; Conservative 0; Mismatches 0; Gaps 0;

QY  5 KAAAVLLPVLLA 17
    | | | | | | | | | |
Db  1 KKCAAVLLPVLLA 13

RESULT 31
AAU00635
ID  AAU00635 standard; peptide; 16 AA.
XX
AC  AAU00635;
XX
XX  07-SEP-2001 (first entry)
XX
DT  07-SEP-2001 (first entry)
XX
XX  Human membrane translocating peptide (MTLP) #4.
XX
DE  Membrane translocating peptide; MTLP; human; intracellular gene delivery;
KW  epithelial cell layer; gastrointestinal tract; circulatory system.
XX
OS  Homo sapiens.
XX
PN  WO200127154-A2.
XX
XX  19-APR-2001.
XX
XX  27-SEP-2000; 2000WO-IB001491.
XX
XX  27-SEP-1999; 99US-0156246P.
XX
XX  (OMAH/) O'MAHONY D J.
XX  (LAMB/) LAMBKIN I J.
XX
XX  O'mahony DJ, Lambkin IJ;
XX
XX  WPI; 2001-300212/31.
XX  N-PSDB; AAS00629.
XX
XX  Compositions for enhancing uptake of e.g. drugs or DNA across a cell
XX  membrane, comprise membrane translocating peptides having specific amino
XX  acid sequences or a derivative, fragment, motif, analog or peptidomimetic
XX  of the peptides.
XX
XX  Claim 2; Page 11; 42pp; English.
XX
XX  The sequence represents a human membrane translocated peptide (MTLP).
XX  MTLPs and their related fragments, motifs, derivatives and analogues are
XX  used for enhancing uptake of a pharmaceutically active agent into a cell,
XX  into or out of an intracellular compartment and across a cell layer (for
XX  example, an epithelial cell layer lining the gastrointestinal tract),
XX  either directly or from a pharmaceutically active agent loaded particle,
XX  into the circulatory system of an animal. This method is useful for
XX  intracellular gene delivery, as a rapid screening method for the
XX  identification of MTLPs which retain the functional activity of a full-
XX  length MTLP, as a cell-based screen for assaying the functional activity
XX  of a MTLP and characterising the properties of a MTLP, for diagnosis of a
XX  pathological disorder (by administration of a MTLP-active agent complex
XX  or MTLP-active particle complex comprising a diagnostic agent) and for
XX  preventing or treating a pathological disorder
XX
SQ  Sequence 16 AA;
    Query Match 61.8%; Score 47; DB 4; Length 16;
    Best Local Similarity 84.6%; Pred. No. 0.74;
    Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY  5 KAAAVLLPVLLA 17
    | | | | | | | | | |
Db  1 KKCAAVLLPVLLA 13

RESULT 32
ADB88778
ID  ADB88778 standard; peptide; 16 AA.
XX
AC  ADB88778;
XX
XX  04-DEC-2003 (first entry)
XX
DE  Membrane translocating peptide #4.
XX
XX  Peyer's patch cell; non-Peyer's patch cell; transcription factor;
KW  upregulated protein; antigen; vaccine delivery; M cell;
KW  membrane translocating peptide.
XX
OS  Unidentified.
XX

```





AC ADK50911;  
 XX 04-NOV-2004 (first entry)  
 XX Human C35 / Kaposi FGF signal sequence MST cell-penetrating peptide 123.  
 DE C35 epitope; cytostatic; vaccine; tumour; breast; bladder carcinoma;  
 XX human; Kaposi FGF signal sequence h region MST; cell-penetrating;  
 KW fibroblast growth factor; membrane-translocating signal.  
 XX Homo sapiens.  
 OS Unidentified.  
 XX WO2003104428-A2.  
 PN 18-DEC-2003.  
 XX 10-JUN-2003; 2003WO-US018252.  
 XX 10-JUN-2002; 2002US-0386738P.  
 PR 11-DEC-2002; 2002US-0432241P.  
 PR 23-APR-2003; 2003US-0464650P.  
 XX (VACC-) VACCINEX INC.  
 PA (UYRP) UNIV ROCHESTER.  
 XX Zauderer M, Evans EE, Borrello MA;  
 PI WPI; 2004-062349/06.  
 XX Novel C35 polypeptide useful for formulation of immunogenic composition  
 PT to induce antibodies and cell-mediated immunity against tumor cells.  
 PT Disclosure; Page 223; 626pp; English.  
 PS The invention relates to a novel isolated polypeptide comprising or  
 CC consisting of two or more C35 peptide epitopes. The polypeptide of the  
 CC invention demonstrates cytostatic activity and may be useful for the  
 CC formulation of an immunogenic composition, such as a vaccine, to induce  
 CC antibodies and cell-mediated immunity against target cells such as tumour  
 CC cells. Furthermore, the polypeptide and its analogues may be useful as  
 CC prognostic markers for carcinoma, such as human breast or bladder  
 CC carcinoma. The current sequence is that of human C35 / Kaposi FGF  
 CC (fibroblast growth factor) signal sequence MST (membrane-translocating  
 CC sequence) cell-penetrating polypeptide of the invention.  
 XX Sequence 21 AA;  
 SQ Query Match 61.8%; Score 47; DB 8; Length 21;  
 Best Local Similarity 100.0%; Pred. No. 0.98;  
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 7 AAVALLPVLLA 17  
 Db 9 AAVALLPVLLA 19  
 RESULT 38  
 ADK50895  
 ID ADK50895 standard; peptide; 21 AA.  
 XX AC ADK50895;  
 XX 04-NOV-2004 (first entry)  
 XX Human C35 / Kaposi FGF signal sequence MST cell-penetrating peptide 107.  
 DE C35 epitope; cytostatic; vaccine; tumour; breast; bladder carcinoma;  
 KW human; Kaposi FGF signal sequence h region MST; cell-penetrating;  
 KW fibroblast growth factor; membrane-translocating signal.  
 XX Homo sapiens.  
 OS Unidentified.

XX WO2003104428-A2.  
 PN 18-DEC-2003.  
 XX 10-JUN-2003; 2003WO-US018252.  
 XX 10-JUN-2002; 2002US-0386738P.  
 PR 11-DEC-2002; 2002US-0432241P.  
 PR 23-APR-2003; 2003US-0464650P.  
 XX (VACC-) VACCINEX INC.  
 PA (UYRP) UNIV ROCHESTER.  
 XX Zauderer M, Evans EE, Borrello MA;  
 PI WPI; 2004-062349/06.  
 XX Novel C35 polypeptide useful for formulation of immunogenic composition  
 PT to induce antibodies and cell-mediated immunity against tumor cells.  
 PT Disclosure; Page 219; 626pp; English.  
 PS The invention relates to a novel isolated polypeptide comprising or  
 CC consisting of two or more C35 peptide epitopes. The polypeptide of the  
 CC invention demonstrates cytostatic activity and may be useful for the  
 CC formulation of an immunogenic composition, such as a vaccine, to induce  
 CC antibodies and cell-mediated immunity against target cells such as tumour  
 CC cells. Furthermore, the polypeptide and its analogues may be useful as  
 CC prognostic markers for carcinoma, such as human breast or bladder  
 CC carcinoma. The current sequence is that of human C35 / Kaposi FGF  
 CC (fibroblast growth factor) signal sequence MST (membrane-translocating  
 CC sequence) cell-penetrating polypeptide of the invention.  
 XX Sequence 21 AA;  
 SQ Query Match 61.8%; Score 47; DB 8; Length 21;  
 Best Local Similarity 91.7%; Pred. No. 0.98;  
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 6 AAVALLPVLLA 17  
 Db 8 AAVALLPVLLA 19  
 RESULT 39  
 ADK50915  
 ID ADK50915 standard; peptide; 21 AA.  
 XX AC ADK50915;  
 XX 04-NOV-2004 (first entry)  
 XX Human C35 / Kaposi FGF signal sequence MST cell-penetrating peptide 127.  
 DE C35 epitope; cytostatic; vaccine; tumour; breast; bladder carcinoma;  
 KW human; Kaposi FGF signal sequence h region MST; cell-penetrating;  
 KW fibroblast growth factor; membrane-translocating signal.  
 XX Homo sapiens.  
 OS Unidentified.  
 XX WO2003104428-A2.  
 PN 18-DEC-2003.  
 XX 10-JUN-2003; 2003WO-US018252.  
 XX 10-JUN-2002; 2002US-0386738P.  
 PR 11-DEC-2002; 2002US-0432241P.  
 PR 23-APR-2003; 2003US-0464650P.  
 XX (VACC-) VACCINEX INC.  
 PA





Query Match 61.8%; Score 47; DB 8; Length 22;  
 Best Local Similarity 91.7%; Pred. No. 1;  
 Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 6 AAAAAVLLPVLLA 17  
 | | | | | | | | | |  
 Db 9 AAAAAVLLPVLLA 20

RESULT 42  
 AAW56392  
 ID AAW56392 standard; peptide; 26 AA.  
 AC AAW56392;  
 DT 05-AUG-1998 (first entry)  
 DE HIV-1MEM polypeptide used to inhibit kappa-Ig light chain expression.

XX SV40MEM polypeptide; signal peptide; fibroblast growth factor;  
 KW SV40 large antigen; nuclear localisation signal; NLS;  
 KW immunosuppressive activity; inhibition; nuclear translocation inhibitor;  
 KW kappa immunoglobulin light chain expression; S. typhosa LPS;  
 KW nuclear translocation; treatment; immune disorder; autoimmune disease;  
 KW hypersensitivity; sepsis; prevention; septic shock; antiviral agent;  
 KW tumour growth suppressor; HIV-1MEM.  
 XX Synthetic.  
 OS  
 XX WO9811907-A1.  
 PN  
 XX 26-MAR-1998.  
 PD  
 XX 15-SEP-1997; 97WO-US016217.  
 PF  
 XX 20-SEP-1996; 96US-0026978P.  
 PR 12-SEP-1997; 97US-00928958.  
 XX (BRIM ) BRISTOL-MYERS SQUIBB CO.  
 PA  
 XX Nadler SG, Cleaveland JS, Blake J, Haffar OK;  
 PI WPI; 1998-217028/19.  
 DR  
 XX Nuclear translocation inhibitor polypeptides - comprising signal sequence  
 PT for delivery through the cytoplasmic membrane and at least 2 nuclear  
 PT localisation sequences.  
 PS Example 1; Page 29; 69pp; English.

CC The present sequence represents the HIV-1MEM polypeptide, which comprises  
 CC the membrane translocation domain of fibroblast growth factor flanked by  
 CC the nuclear localisation signal (NLS) from the HIV-1 matrix protein. The  
 CC immunosuppressive activity of this peptide was compared with that of the  
 CC SV40MEM polypeptide (see AAW56391). The SV40MEM polypeptide causes  
 CC approximately 75-80% inhibition of kappa immunoglobulin (Ig) light chain  
 CC expression in response to S. typhosa LPS. L- and D-forms of the SV40MEM  
 CC peptide are equally effective. The SV40MEM polypeptide exemplifies the  
 CC nuclear translocation inhibitor polypeptide of the invention. Nuclear  
 CC translocation inhibitor polypeptides comprise a signal sequence peptide  
 CC capable of delivering the polypeptide through the cytoplasmic membrane  
 CC into a cell, and at least 2 nuclear localisation sequences (NLSs). The  
 CC polypeptides can be used to inhibit nuclear translocation of a cellular  
 CC protein. In addition, since the nuclear translocation of certain cellular  
 CC peptides is required for the host organism to mount an immune response,  
 CC the polypeptide inhibitors are useful as immunosuppression agents. The  
 CC polypeptides can therefore be used for the treatment of immune disorders  
 CC including autoimmune diseases. The polypeptides can also be used for  
 CC treating physical symptoms manifested by responses to allergens which can  
 CC initiate a state of hypersensitivity, for the treatment of sepsis and in  
 CC the prevention of septic shock, antiviral agents, tumour growth  
 CC suppressors, and for transcriptionally modulating the expression of  
 CC cellular genes

XX SQ Sequence 26 AA;  
 Query Match 61.8%; Score 47; DB 2; Length 26;  
 Best Local Similarity 70.6%; Pred. No. 1.2;  
 Matches 12; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 KKKRKAATAVLLPVLLA 17  
 | | | | | | | | | |  
 Db 1 KKKYKAATAVLLPVLLA 17

RESULT 43  
 AAY55813  
 ID AAY55813 standard; peptide; 26 AA.  
 XX  
 AC AAY55813;  
 DT 06-MAR-2000 (first entry)  
 DE HIV-1MEM polypeptide fragment.

XX Cellular protein; nuclear translocation; nuclear localization signal;  
 KW immunosuppressant; immune response; viral infection; immune disorder;  
 KW rheumatoid arthritis; multiple sclerosis; juvenile-onset diabetes; NLS;  
 KW systemic lupus erythematosus; SLE; asthma; sepsis; tumor growth; HIV;  
 KW fibroblast growth factor; matrix protein.  
 XX Synthetic.  
 OS Human immunodeficiency virus 1.  
 XX WO9957138-A1.  
 PN  
 XX 11-NOV-1999.  
 PD  
 XX 26-APR-1999; 99WO-US008984.  
 PF  
 XX 04-MAY-1998; 98US-00072429.  
 PR  
 XX (BRIM ) BRISTOL-MYERS SQUIBB CO.  
 PA  
 XX Nadler SG;  
 PI WPI; 2000-062141/05.  
 DR  
 XX Novel composition comprising a nuclear localization signal and an  
 PT inhibitor of nuclear translocation, useful for preventing transplanted  
 PT organ or tissue rejection.  
 PS Example 1; Page 26; 68pp; English.

CC The invention provides a composition comprising an isolated polypeptide  
 CC (I) (polypeptide inhibitor of a cellular protein nuclear translocation)  
 CC which comprises a signal sequence peptide capable of delivering (I)  
 CC through a cytoplasmic membrane into the cell; and a nuclear localization  
 CC signal sequence (NLS) present along with an immunosuppressant (I) or its  
 CC derivatives provide useful tools for introducing an exogenous polypeptide  
 CC comprising an NLS into an intact cell to inhibit nuclear translocation of  
 CC a cellular protein, for studying the role of nuclear translocation in the  
 CC regulation of cellular processes. See AAY55812 for detailed uses of (I)  
 CC and compositions containing (I). The present sequence represents a HIV-  
 CC 1MEM polypeptide fragment comprising the membrane translocation domain of  
 CC fibroblast growth factor flanked by NLS from HIV-1 matrix protein

XX SQ Sequence 26 AA;  
 Query Match 61.8%; Score 47; DB 3; Length 26;  
 Best Local Similarity 70.6%; Pred. No. 1.2;  
 Matches 12; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 KKKRKAATAVLLPVLLA 17  
 | | | | | | | | | |  
 Db 1 KKKYKAATAVLLPVLLA 17

```

RESULT 44
ADK50900
ID ADK50900 standard; peptide; 33 AA.
XX AC ADK50900;
XX DT 04-NOV-2004 (first entry)
XX DE Human C35 / Kaposi FGF signal sequence MST cell-penetrating peptide 112.
XX KW C35 epitope; cytostatic; vaccine; tumour; breast; bladder carcinoma;
XX KW human; Kaposi FGF signal sequence h region MST; cell-penetrating;
XX KW fibroblast growth factor; membrane-translocating signal.
XX OS Homo sapiens.
XX OS Unidentified.
XX PN WO2003104428-A2.
XX PD 18-DEC-2003.
XX PF 10-JUN-2003; 2003WO-US018252.
XX PR 10-JUN-2002; 2002US-0386738P.
XX PR 11-DEC-2002; 2002US-0432241P.
XX PR 23-APR-2003; 2003US-0464650P.
XX PA (VACC-) VACCINEX INC.
XX PA (UYRP ) UNIV ROCHESTER.
XX PI Zauderer M, Evans EE, Borrello MA;
XX DR WPI; 2004-062349/06.
XX PF 10-JUN-2003; 2003WO-US018252.
XX PR 10-JUN-2002; 2002US-0386738P.
XX PR 11-DEC-2002; 2002US-0432241P.
XX PR 23-APR-2003; 2003US-0464650P.
XX PA (VACC-) VACCINEX INC.
XX PA (UYRP ) UNIV ROCHESTER.
XX PI Zauderer M, Evans EE, Borrello MA;
XX DR WPI; 2004-062349/06.
XX KW Novel C35 polypeptide useful for formulation of immunogenic composition
XX PT to induce antibodies and cell-mediated immunity against tumor cells.
XX PS Disclosure; Page 220; 626pp; English.
XX CC The invention relates to a novel isolated polypeptide comprising or
XX CC consisting of two or more C35 peptide epitopes. The polypeptide of the
XX CC invention demonstrates cytostatic activity and may be useful for the
XX CC formulation of an immunogenic composition, such as a vaccine, to induce
XX CC antibodies and cell-mediated immunity against target cells such as tumor
XX CC cells. Furthermore, the polypeptide and its analogues may be useful as
XX CC prognostic markers for carcinoma, such as human breast or bladder
XX CC carcinoma. The current sequence is that of human C35 / Kaposi FGF
XX CC (fibroblast growth factor) signal sequence MST (membrane-translocating
XX CC sequence) cell-penetrating polypeptide of the invention.
XX SQ Sequence 33 AA;
XX Query Match 61.8%; Score 47; DB 8; Length 33;
XX Best Local Similarity 91.7%; Pred. No. 1.6;
XX Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX QY 6 AAAAALLPVLLA 17
XX DB 20 AVAAVLLPVLLA 31
XX DE
XX DE Human C35 / Kaposi FGF signal sequence MST cell-penetrating peptide 128.
XX KW C35 epitope; cytostatic; vaccine; tumour; breast; bladder carcinoma;
XX KW human; Kaposi FGF signal sequence h region MST; cell-penetrating;
XX KW fibroblast growth factor; membrane-translocating signal.
XX OS Homo sapiens.
XX OS Unidentified.
XX PN WO2003104428-A2.
XX PD 18-DEC-2003.
XX PF 10-JUN-2003; 2003WO-US018252.
XX PR 10-JUN-2002; 2002US-0386738P.
XX PR 11-DEC-2002; 2002US-0432241P.
XX PR 23-APR-2003; 2003US-0464650P.
XX PA (VACC-) VACCINEX INC.
XX PA (UYRP ) UNIV ROCHESTER.
XX PI Zauderer M, Evans EE, Borrello MA;
XX DR WPI; 2004-062349/06.
XX KW Novel C35 polypeptide useful for formulation of immunogenic composition
XX PT to induce antibodies and cell-mediated immunity against tumor cells.
XX PS Disclosure; Page 220; 626pp; English.
XX CC The invention relates to a novel isolated polypeptide comprising or
XX CC consisting of two or more C35 peptide epitopes. The polypeptide of the
XX CC invention demonstrates cytostatic activity and may be useful for the
XX CC formulation of an immunogenic composition, such as a vaccine, to induce
XX CC antibodies and cell-mediated immunity against target cells such as tumor
XX CC cells. Furthermore, the polypeptide and its analogues may be useful as
XX CC prognostic markers for carcinoma, such as human breast or bladder
XX CC carcinoma. The current sequence is that of human C35 / Kaposi FGF
XX CC (fibroblast growth factor) signal sequence MST (membrane-translocating
XX CC sequence) cell-penetrating polypeptide of the invention.
XX SQ Sequence 33 AA;
XX Query Match 61.8%; Score 47; DB 8; Length 33;
XX Best Local Similarity 91.7%; Pred. No. 1.6;
XX Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX QY 6 AAAAALLPVLLA 17
XX DB 20 AVAAVLLPVLLA 31
XX DE
XX DE Human C35 / Kaposi FGF signal sequence MST cell-penetrating peptide 20.
XX KW C35 epitope; cytostatic; vaccine; tumour; breast; bladder carcinoma;
XX KW human; Kaposi FGF signal sequence h region MST; cell-penetrating;
XX KW fibroblast growth factor; membrane-translocating signal.
XX OS Homo sapiens.
XX OS Unidentified.
XX PN WO2003104428-A2.
XX PD 18-DEC-2003.
XX PF 10-JUN-2003; 2003WO-US018252.

```

```

KW C35 epitope; cytostatic; vaccine; tumour; breast; bladder carcinoma;
KW human; Kaposi FGF signal sequence h region MST; cell-penetrating;
KW fibroblast growth factor; membrane-translocating signal.
XX OS Homo sapiens.
XX OS Unidentified.
XX PN WO2003104428-A2.
XX PD 18-DEC-2003.
XX PF 10-JUN-2003; 2003WO-US018252.
XX PR 10-JUN-2002; 2002US-0386738P.
XX PR 11-DEC-2002; 2002US-0432241P.
XX PR 23-APR-2003; 2003US-0464650P.
XX PA (VACC-) VACCINEX INC.
XX PA (UYRP ) UNIV ROCHESTER.
XX PI Zauderer M, Evans EE, Borrello MA;
XX DR WPI; 2004-062349/06.
XX PF Novel C35 polypeptide useful for formulation of immunogenic composition
XX PT to induce antibodies and cell-mediated immunity against tumor cells.
XX PS Disclosure; Page 197; 626pp; English.
XX CC The invention relates to a novel isolated polypeptide comprising or
XX CC consisting of two or more C35 peptide epitopes. The polypeptide of the
XX CC invention demonstrates cytostatic activity and may be useful for the
XX CC formulation of an immunogenic composition, such as a vaccine, to induce
XX CC antibodies and cell-mediated immunity against target cells such as tumor
XX CC cells. Furthermore, the polypeptide and its analogues may be useful as
XX CC prognostic markers for carcinoma, such as human breast or bladder
XX CC carcinoma. The current sequence is that of human C35 / Kaposi FGF
XX CC (fibroblast growth factor) signal sequence MST (membrane-translocating
XX CC sequence) cell-penetrating polypeptide of the invention.
XX SQ Sequence 33 AA;
XX Query Match 61.8%; Score 47; DB 8; Length 33;
XX Best Local Similarity 91.7%; Pred. No. 1.6;
XX Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX QY 6 AAAAALLPVLLA 17
XX DB 20 AVAAVLLPVLLA 31
XX DE
XX DE Human C35 / Kaposi FGF signal sequence MST cell-penetrating peptide 128.
XX KW C35 epitope; cytostatic; vaccine; tumour; breast; bladder carcinoma;
XX KW human; Kaposi FGF signal sequence h region MST; cell-penetrating;
XX KW fibroblast growth factor; membrane-translocating signal.
XX OS Homo sapiens.
XX OS Unidentified.
XX PN WO2003104428-A2.
XX PD 18-DEC-2003.
XX PF 10-JUN-2003; 2003WO-US018252.

```

XX  
PR 10-JUN-2002; 2002US-0386738P.  
PR 11-DEC-2002; 2002US-0432241P.  
PR 23-APR-2003; 2003US-0464650P.  
XX  
XX (VACC-) VACCINEX INC.  
XX (UYRP) UNIV ROCHESTER.  
XX  
XX Zauderer M, Evans EE, Borrello MA;  
XX WPI; 2004-062349/06.  
XX  
XX Novel C35 polypeptide useful for formulation of immunogenic composition  
PT to induce antibodies and cell-mediated immunity against tumor cells.  
XX  
XX Disclosure; Page 224; 626pp; English.  
XX  
XX The invention relates to a novel isolated polypeptide comprising or  
CC consisting of two or more C35 peptide epitopes. The polypeptide of the  
CC invention demonstrates cytostatic activity and may be useful for the  
CC formulation of an immunogenic composition, such as a vaccine, to induce  
CC antibodies and cell-mediated immunity against target cells such as tumor  
CC cells. Furthermore, the polypeptide and its analogues may be useful as  
CC prognostic markers for carcinoma, such as human breast or bladder  
CC carcinoma. The current sequence is that of human C35 / Kaposi FGF  
CC (fibroblast growth factor) signal sequence MST (membrane-translocating  
CC sequence) cell-penetrating polypeptide of the invention.  
XX  
XX Sequence 33 AA;  
XX  
XX The invention relates to a novel isolated polypeptide comprising or  
CC consisting of two or more C35 peptide epitopes. The polypeptide of the  
CC invention demonstrates cytostatic activity and may be useful for the  
CC formulation of an immunogenic composition, such as a vaccine, to induce  
CC antibodies and cell-mediated immunity against target cells such as tumor  
CC cells. Furthermore, the polypeptide and its analogues may be useful as  
CC prognostic markers for carcinoma, such as human breast or bladder  
CC carcinoma. The current sequence is that of human C35 / Kaposi FGF  
CC (fibroblast growth factor) signal sequence MST (membrane-translocating  
CC sequence) cell-penetrating polypeptide of the invention.  
XX  
XX Sequence 33 AA;  
XX  
XX Query Match 61.8%; Score 47; DB 8; Length 33;  
Best Local Similarity 100.0%; Pred. No. 1.6;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
XX  
XX QY 7 ARAVLLPVLLA 17  
DB 21 ARAVLLPVLLA 31  
XX  
XX RESULT 47  
ADKS0896  
XX ID ADKS0896 standard; peptide; 33 AA.  
XX AC  
XX ADKS0896;  
XX  
XX 04-NOV-2004 (first entry)  
XX  
XX Human C35 / Kaposi FGF signal sequence MST cell-penetrating peptide 108.  
XX  
XX C35 epitope; cytostatic; vaccine; tumour; breast; bladder carcinoma;  
KW human; Kaposi FGF signal sequence h region MST; cell-penetrating;  
KW fibroblast growth factor; membrane-translocating signal.  
XX  
XX Homo sapiens.  
OS Unidentified.  
XX WO2003104428-A2.  
XX  
XX 18-DEC-2003.  
XX  
XX 10-JUN-2003; 2003WO-US018252.  
XX  
XX C35 epitope; cytostatic; vaccine; tumour; breast; bladder carcinoma;  
KW human; Kaposi FGF signal sequence h region MST; cell-penetrating;  
KW fibroblast growth factor; membrane-translocating signal.  
XX  
XX Homo sapiens.  
OS Unidentified.  
XX WO2003104428-A2.  
XX  
XX 18-DEC-2003.  
XX  
XX 10-JUN-2003; 2003WO-US018252.  
XX  
XX 10-JUN-2002; 2002US-0386738P.  
PR 11-DEC-2002; 2002US-0432241P.  
PR 23-APR-2003; 2003US-0464650P.  
XX  
XX (VACC-) VACCINEX INC.  
XX (UYRP) UNIV ROCHESTER.  
XX  
XX Zauderer M, Evans EE, Borrello MA;  
XX WPI; 2004-062349/06.  
XX  
XX Novel C35 polypeptide useful for formulation of immunogenic composition  
PT to induce antibodies and cell-mediated immunity against tumor cells.  
XX  
XX Disclosure; Page 224; 626pp; English.  
XX  
XX The invention relates to a novel isolated polypeptide comprising or  
CC consisting of two or more C35 peptide epitopes. The polypeptide of the  
CC invention demonstrates cytostatic activity and may be useful for the  
CC formulation of an immunogenic composition, such as a vaccine, to induce  
CC antibodies and cell-mediated immunity against target cells such as tumor  
CC cells. Furthermore, the polypeptide and its analogues may be useful as  
CC prognostic markers for carcinoma, such as human breast or bladder  
CC carcinoma. The current sequence is that of human C35 / Kaposi FGF  
CC (fibroblast growth factor) signal sequence MST (membrane-translocating  
CC sequence) cell-penetrating polypeptide of the invention.  
XX  
XX Sequence 33 AA;  
XX  
XX Query Match 61.8%; Score 47; DB 8; Length 33;  
Best Local Similarity 100.0%; Pred. No. 1.6;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
XX  
XX QY 7 ARAVLLPVLLA 17  
DB 21 ARAVLLPVLLA 31  
XX  
XX RESULT 47  
ADKS0896  
XX ID ADKS0896 standard; peptide; 33 AA.  
XX AC  
XX ADKS0896;  
XX  
XX 04-NOV-2004 (first entry)  
XX  
XX Human C35 / Kaposi FGF signal sequence MST cell-penetrating peptide 124.  
XX  
XX C35 epitope; cytostatic; vaccine; tumour; breast; bladder carcinoma;  
KW human; Kaposi FGF signal sequence h region MST; cell-penetrating;  
KW fibroblast growth factor; membrane-translocating signal.  
XX  
XX Homo sapiens.  
OS Unidentified.  
XX WO2003104428-A2.  
XX  
XX 18-DEC-2003.  
XX  
XX 10-JUN-2003; 2003WO-US018252.  
XX  
XX C35 epitope; cytostatic; vaccine; tumour; breast; bladder carcinoma;  
KW human; Kaposi FGF signal sequence h region MST; cell-penetrating;  
KW fibroblast growth factor; membrane-translocating signal.  
XX  
XX Homo sapiens.  
OS Unidentified.  
XX WO2003104428-A2.  
XX  
XX 18-DEC-2003.  
XX  
XX 10-JUN-2003; 2003WO-US018252.  
XX  
XX 10-JUN-2002; 2002US-0386738P.  
PR 11-DEC-2002; 2002US-0432241P.  
PR 23-APR-2003; 2003US-0464650P.  
XX  
XX (VACC-) VACCINEX INC.  
XX (UYRP) UNIV ROCHESTER.  
XX  
XX Zauderer M, Evans EE, Borrello MA;  
XX WPI; 2004-062349/06.  
XX  
XX Novel C35 polypeptide useful for formulation of immunogenic composition  
PT to induce antibodies and cell-mediated immunity against tumor cells.  
XX  
XX Disclosure; Page 223; 626pp; English.  
XX  
XX The invention relates to a novel isolated polypeptide comprising or  
CC consisting of two or more C35 peptide epitopes. The polypeptide of the  
CC invention demonstrates cytostatic activity and may be useful for the  
CC formulation of an immunogenic composition, such as a vaccine, to induce  
CC antibodies and cell-mediated immunity against target cells such as tumor  
CC cells. Furthermore, the polypeptide and its analogues may be useful as  
CC prognostic markers for carcinoma, such as human breast or bladder  
CC carcinoma. The current sequence is that of human C35 / Kaposi FGF  
CC (fibroblast growth factor) signal sequence MST (membrane-translocating  
CC sequence) cell-penetrating polypeptide of the invention.  
XX  
XX Sequence 33 AA;  
XX  
XX Query Match 61.8%; Score 47; DB 8; Length 33;  
Best Local Similarity 91.7%; Pred. No. 1.6;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
XX  
XX QY 6 ARAVLLPVLLA 17  
DB 20 ARAVLLPVLLA 31  
XX  
XX RESULT 48  
ADKS0912  
XX ID ADKS0912 standard; peptide; 33 AA.  
XX AC  
XX ADKS0912;  
XX  
XX 04-NOV-2004 (first entry)  
XX  
XX Human C35 / Kaposi FGF signal sequence MST cell-penetrating peptide 124.  
XX  
XX C35 epitope; cytostatic; vaccine; tumour; breast; bladder carcinoma;  
KW human; Kaposi FGF signal sequence h region MST; cell-penetrating;  
KW fibroblast growth factor; membrane-translocating signal.  
XX  
XX Homo sapiens.  
OS Unidentified.  
XX WO2003104428-A2.  
XX  
XX 18-DEC-2003.  
XX  
XX 10-JUN-2003; 2003WO-US018252.  
XX  
XX 10-JUN-2002; 2002US-0386738P.  
PR 11-DEC-2002; 2002US-0432241P.  
PR 23-APR-2003; 2003US-0464650P.  
XX  
XX (VACC-) VACCINEX INC.  
XX (UYRP) UNIV ROCHESTER.  
XX  
XX Zauderer M, Evans EE, Borrello MA;  
XX WPI; 2004-062349/06.  
XX  
XX Novel C35 polypeptide useful for formulation of immunogenic composition  
PT to induce antibodies and cell-mediated immunity against tumor cells.  
XX  
XX Disclosure; Page 223; 626pp; English.  
XX  
XX The invention relates to a novel isolated polypeptide comprising or  
CC consisting of two or more C35 peptide epitopes. The polypeptide of the  
CC invention demonstrates cytostatic activity and may be useful for the  
CC formulation of an immunogenic composition, such as a vaccine, to induce  
CC antibodies and cell-mediated immunity against target cells such as tumor  
CC cells. Furthermore, the polypeptide and its analogues may be useful as  
CC prognostic markers for carcinoma, such as human breast or bladder  
CC carcinoma. The current sequence is that of human C35 / Kaposi FGF  
CC (fibroblast growth factor) signal sequence MST (membrane-translocating  
CC sequence) cell-penetrating polypeptide of the invention.  
XX  
XX Sequence 33 AA;  
XX  
XX Query Match 61.8%; Score 47; DB 8; Length 33;  
Best Local Similarity 91.7%; Pred. No. 1.6;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
XX  
XX QY 6 ARAVLLPVLLA 17  
DB 20 ARAVLLPVLLA 31  
XX  
XX RESULT 48  
ADKS0912  
XX ID ADKS0912 standard; peptide; 33 AA.  
XX AC  
XX ADKS0912;  
XX  
XX 04-NOV-2004 (first entry)  
XX  
XX Human C35 / Kaposi FGF signal sequence MST cell-penetrating peptide 124.  
XX  
XX C35 epitope; cytostatic; vaccine; tumour; breast; bladder carcinoma;  
KW human; Kaposi FGF signal sequence h region MST; cell-penetrating;  
KW fibroblast growth factor; membrane-translocating signal.  
XX  
XX Homo sapiens.  
OS Unidentified.  
XX WO2003104428-A2.  
XX  
XX 18-DEC-2003.  
XX  
XX 10-JUN-2003; 2003WO-US018252.  
XX  
XX 10-JUN-2002; 2002US-0386738P.  
PR 11-DEC-2002; 2002US-0432241P.  
PR 23-APR-2003; 2003US-0464650P.  
XX  
XX (VACC-) VACCINEX INC.  
XX (UYRP) UNIV ROCHESTER.  
XX  
XX Zauderer M, Evans EE, Borrello MA;  
XX WPI; 2004-062349/06.  
XX  
XX Novel C35 polypeptide useful for formulation of immunogenic composition  
PT to induce antibodies and cell-mediated immunity against tumor cells.  
XX  
XX Disclosure; Page 223; 626pp; English.  
XX  
XX The invention relates to a novel isolated polypeptide comprising or  
CC consisting of two or more C35 peptide epitopes. The polypeptide of the  
CC invention demonstrates cytostatic activity and may be useful for the  
CC formulation of an immunogenic composition, such as a vaccine, to induce  
CC antibodies and cell-mediated immunity against target cells such as tumor  
CC cells. Furthermore, the polypeptide and its analogues may be useful as  
CC prognostic markers for carcinoma, such as human breast or bladder  
CC carcinoma. The current sequence is that of human C35 / Kaposi FGF  
CC (fibroblast growth factor) signal sequence MST (membrane-translocating  
CC sequence) cell-penetrating polypeptide of the invention.  
XX  
XX Sequence 33 AA;  
XX  
XX Query Match 61.8%; Score 47; DB 8; Length 33;  
Best Local Similarity 91.7%; Pred. No. 1.6;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
XX  
XX QY 6 ARAVLLPVLLA 17  
DB 20 ARAVLLPVLLA 31  
XX  
XX

CC carcinoma. The current sequence is that of human C35 / Kaposi FGF  
CC (fibroblast growth factor) signal sequence MST (membrane-translocating  
CC sequence) cell-penetrating polypeptide of the invention.  
XX  
SQ Sequence 33 AA;

Query Match 61.8%; Score 47; DB 8; Length 33;  
Best Local Similarity 100.0%; Pred. No. 1.6;  
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 AAVALLPVLLA 17  
|||||  
Db 21 AAVALLPVLLA 31

RESULT 49  
ADK50924  
ID ADK50924 standard; peptide; 34 AA.  
XX  
AC ADK50924;  
XX  
DT 04-NOV-2004 (first entry)  
XX  
DE Human C35 / Kaposi FGF signal sequence MST cell-penetrating peptide 136.  
XX  
KW C35 epitope; cytostatic; vaccine; tumour; breast; bladder carcinoma;  
KW human; Kaposi FGF signal sequence h region MST; cell-penetrating;  
KW fibroblast growth factor; membrane-translocating signal.  
XX  
OS Homo sapiens.  
OS Unidentified.  
XX  
PN W02003104428-A2.  
XX  
PD 18-DEC-2003.  
XX  
PF 10-JUN-2003; 2003WO-US018252.  
XX  
PR 10-JUN-2002; 2002US-0386738P.  
PR 11-DEC-2002; 2002US-0432241P.  
PR 23-APR-2003; 2003US-0464650P.  
XX  
PA (VACC-) VACCINEX INC.  
PA (UYRP) UNIV ROCHESTER.  
XX  
PI Zauderer M, Evans ES, Borrello MA;  
XX  
DR WPI; 2004-062349/06.  
XX  
PT Novel C35 polypeptide useful for formulation of immunogenic composition  
PT to induce antibodies and cell-mediated immunity against tumor cells.  
XX  
PS Disclosure; Page 226; 626pp; English.  
XX

The invention relates to a novel isolated polypeptide comprising or  
consisting of two or more C35 peptide epitopes. The polypeptide of the  
invention demonstrates cytostatic activity and may be useful for the  
formulation of an immunogenic composition, such as a vaccine, to induce  
antibodies and cell-mediated immunity against target cells such as tumour  
cells. Furthermore, the polypeptide and its analogues may be useful as  
prognostic markers for carcinoma, such as human breast or bladder  
carcinoma. The current sequence is that of human C35 / Kaposi FGF  
CC (fibroblast growth factor) signal sequence MST (membrane-translocating  
CC sequence) cell-penetrating polypeptide of the invention.  
XX  
SQ Sequence 34 AA;

Query Match 61.8%; Score 47; DB 8; Length 34;  
Best Local Similarity 91.7%; Pred. No. 1.6;  
Matches 11; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 6 AAVALLPVLLA 17  
|||||

Db 21 AAVALLPVLLA 32

RESULT 50  
AAW56414  
ID AAW56414 standard; peptide; 29 AA.  
XX  
AC AAW56414;  
XX  
DT 05-AUG-1998 (first entry)  
XX  
DE Nuclear translocation inhibitor polypeptide of the invention.  
XX  
KW Signal peptide; nuclear localisation signal; NLS;  
KW immunosuppressive activity; inhibition; nuclear translocation inhibitor;  
KW nuclear translocation; treatment; immune disorder; autoimmune disease;  
KW hypersensitivity; sepsis; prevention; septic shock; antiviral agent;  
KW tumour growth suppressor.  
XX  
OS Unidentified.  
XX  
PN W09811907-A1.  
XX  
PD 26-MAR-1998.  
XX  
PF 15-SEP-1997; 97WO-US016217.  
XX  
PR 20-SEP-1996; 96US-0026978P.  
PR 12-SEP-1997; 97US-00928958.  
XX  
PA (BRIM) BRISTOL-MYERS SQUIBB CO.  
XX  
PI Nadler SG, Cleaveland JS, Blake J, Haffar OK;  
XX  
DR WPI; 1998-217028/19.  
XX  
PT Nuclear translocation inhibitor polypeptides - comprising signal sequence  
PT for delivery through the cytoplasmic membrane and at least 2 nuclear  
PT localisation sequences.  
XX  
PS Claim 14; Page 45; 69pp; English.  
XX

The present sequence exemplifies the nuclear translocation inhibitor  
polypeptide of the invention. Nuclear translocation inhibitor  
polypeptides comprise a signal sequence peptide capable of delivering the  
polypeptide through the cytoplasmic membrane into a cell, and at least 2  
nuclear localisation signals (NLSs). The polypeptides can be used to  
inhibit nuclear translocation of a cellular protein. In addition, since  
the nuclear translocation of certain cellular peptides is required for  
the host organism to mount an immune response, the polypeptide inhibitors  
are useful as immunosuppression agents. The polypeptides can therefore be  
used for the treatment of immune disorders including autoimmune diseases.  
CC The polypeptides can also be used for treating physical symptoms  
CC manifested by responses to allergens which can initiate a state of  
CC hypersensitivity, for the treatment of sepsis and in the prevention of  
CC septic shock, antiviral agents, tumour growth suppressors, and for  
CC transcriptionally modulating the expression of cellular genes  
XX  
SQ Sequence 29 AA;

Query Match 61.2%; Score 46.5; DB 2; Length 29;  
Best Local Similarity 72.2%; Pred. No. 1.7;  
Matches 13; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

QY 1 KKXKRAAAVLLP-VLLA 17  
|||||  
Db 2 KKXRVVAVALP-VLLA 19

Search completed: June 2, 2005, 01:35:42  
Job time : 141.312 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 2, 2005, 01:26:26 ; Search time 120.594 Seconds  
(without alignments)  
72.187 Million cell updates/sec

Title: US-10-764-235-24

Perfect score: 76  
Sequence: 1 KKKRKAARVLLPVLLA 17

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 750 summaries

Database : UniProt\_03.\*

1: uniprot\_sprot.\*  
2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	45	59.2	420	2 Q6CK41	Q6CK41 kluyveromyc
2	44	57.9	157	2 Q7SBW3	Q7SBW3 neurospora
3	44	57.9	242	2 Q8LR32	Q8LR32 oryza sativ
4	44	57.9	802	2 Q8LAU4	Q8LAU4 oryza sativ
5	43	56.6	519	1 PPBT_CHICK	Q92058 gallus gall
6	43	56.6	582	2 Q99PG1	Q99PG1 cricetus
7	42	55.3	133	2 Q6ND81	Q6ND81 rhodospaudo
8	42	55.3	210	2 Q9CR24	Q9CR24 m mus muscu
9	42	55.3	221	2 Q72PH8	Q72PH8 leptospira
10	42	55.3	221	2 Q8F6U9	Q8F6U9 leptospira
11	42	55.3	268	2 Q8G6L8	Q8G6L8 bifidobacte
12	42	55.3	293	2 Q8FYJ9	Q8FYJ9 bruceella su
13	42	55.3	297	2 Q8XJ99	Q8XJ99 bruceella me
14	42	55.3	524	2 Q8K0I2	Q8K0I2 mus musculu
15	42	55.3	527	1 PTIB_BACSU	P54715 bacillus su
16	42	55.3	607	2 Q8BT03	Q8BT03 mus musculu
17	42	55.3	621	2 Q8C4D2	Q8C4D2 mus musculu
18	42	55.3	628	2 Q89IJ7	Q89IJ7 bradyrhizob
19	42	55.3	741	2 Q8BG19	Q8BG19 m mus muscu
20	42	55.3	1049	2 Q78899	Q78899 neurospora
21	41.5	54.6	346	2 Q8S4W4	Q8S4W4 porphyra ye
22	41	53.9	67	2 Q6GX97	Q6GX97 mus musculu
23	41	53.9	216	1 EVGI_MOUSE	Q6D9S1 mus musculu
24	41	53.9	216	2 Q6AYH8	Q6AYH8 rattus norv
25	41	53.9	260	2 Q92M69	Q92M69 rhizobium m
26	41	53.9	262	2 Q9C4D8	Q9C4D8 thraustochy
27	41	53.9	297	2 Q7OB91	Q7OB91 anopheles g
28	41	53.9	342	2 Q5Z264	Q5Z264 treponema d
29	41	53.9	346	2 Q7JKN6	Q7JKN6 treponema d
30	41	53.9	403	2 Q89S27	Q89S27 bradyrhizob
31	41	53.9	492	2 Q8PTS8	Q8PTS8 corynebacte

32	41	53.9	524	2 Q9BGZ6	Q9BGZ6 macaca fasc
33	41	53.9	528	2 Q25230	Q25230 helicobacte
34	41	53.9	529	2 Q65MC3	Q65MC3 bacillus li
35	41	53.9	530	2 Q62XR6	Q62XR6 bacillus li
36	41	53.9	580	1 MM14_PIG	MM14_PIG sus scrofa
37	41	53.9	582	1 MM14_HUMAN	MM14_HUMAN homo sapien
38	41	53.9	582	1 MM14_MOUSE	MM14_MOUSE mus musculu
39	41	53.9	582	1 MM14_RABIT	MM14_RABIT oryctolagus
40	41	53.9	582	1 MM14_RAT	MM14_RAT rattus norv
41	41	53.9	582	2 Q9GLE4	Q9GLE4 bos taurus
42	41	53.9	582	2 Q9XSP0	Q9XSP0 capra hircu
43	41	53.9	582	2 Q8DFU5	Q8DFU5 mus musculu
44	41	53.9	582	2 Q8BTX2	Q8BTX2 mus musculu
45	41	53.9	886	2 Q6EQK1	Q6EQK1 oryza sativ
46	40.5	53.3	177	2 Q7W3J1	Q7W3J1 bordetella
47	40.5	53.3	177	2 Q7W6W6	Q7W6W6 bordetella
48	40.5	53.3	288	2 Q9VNM7	Q9VNM7 drosophila
49	40.5	53.3	497	2 Q8JND4	Q8JND4 white spot
50	40.5	53.3	544	2 Q9QY99	Q9QY99 white spot
51	40.5	53.3	544	2 Q91L70	Q91L70 white spot
52	40	52.6	264	2 Q8S3Y3	Q8S3Y3 oryza sativ
53	40	52.6	303	2 Q83H01	Q83H01 tropheryma
54	40	52.6	303	2 Q83ID6	Q83ID6 tropheryma
55	40	52.6	306	2 Q8A5G0	Q8A5G0 bacteroides
56	40	52.6	310	2 Q9XX39	Q9XX39 caenorhabdi
57	40	52.6	387	2 Q8TES1	Q8TES1 xyliella fas
58	40	52.6	389	2 Q9PGL7	Q9PGL7 xyliella fas
59	40	52.6	495	2 Q8FTV5	Q8FTV5 corynebacte
60	40	52.6	497	2 Q8N2W2	Q8N2W2 rhodospaudo
61	40	52.6	528	2 Q968M1	Q968M1 trypanosoma
62	40	52.6	529	2 Q8CNAB	Q8CNAB staphylococ
63	40	52.6	534	2 Q99RV0	Q99RV0 staphylococ
64	40	52.6	534	2 Q7A049	Q7A049 staphylococ
65	40	52.6	534	2 Q7A400	Q7A400 staphylococ
66	40	52.6	534	2 Q6G6Z7	Q6G6Z7 staphylococ
67	40	52.6	534	2 Q6GEB1	Q6GEB1 staphylococ
68	40	52.6	563	2 Q802X9	Q802X9 brachydanio
69	40	52.6	567	2 Q73306	Q73306 bacillus ce
70	40	52.6	568	2 Q73306	Q73306 bacillus ce
71	40	52.6	570	2 Q6HFK9	Q6HFK9 bacillus th
72	40	52.6	571	2 Q758H5	Q758H5 ashbya goss
73	40	52.6	644	2 Q6MGU8	Q6MGU8 dellavibri
74	40	52.6	863	1 IMB1_SCHPO	IMB1_SCHPO schizosacch
75	40	52.6	920	2 Q02531	Q02531 schizosacch
76	40	52.6	980	2 Q9ZUI0	Q9ZUI0 arabidopsis
77	40	52.6	999	2 Q7XD3	Q7XD3 oryza sativ
78	40	52.6	1673	2 Q750X6	Q750X6 ashbya goss
79	39.5	52.0	343	2 Q6HFY6	Q6HFY6 bacillus th
80	39	51.3	74	2 Q7V8P7	Q7V8P7 prochloroco
81	39	51.3	119	2 Q6L4S7	Q6L4S7 oryza sativ
82	39	51.3	162	2 Q9HNE5	Q9HNE5 halobacteri
83	39	51.3	165	2 Q8PAS5	Q8PAS5 xanthomonas
84	39	51.3	222	2 Q8CPR9	Q8CPR9 kluyveromyc
85	39	51.3	231	2 Q89UD1	Q89UD1 bradyrhizob
86	39	51.3	245	2 Q72QX0	Q72QX0 leptospira
87	39	51.3	245	2 Q8F4X1	Q8F4X1 leptospira
88	39	51.3	255	2 Q7UIE9	Q7UIE9 rhodopirell
89	39	51.3	257	2 Q9KZC0	Q9KZC0 streptomyce
90	39	51.3	266	2 Q6TM89	Q6TM89 bacterioph
91	39	51.3	303	2 Q9SXY2	Q9SXY2 oryza sativ
92	39	51.3	304	2 Q49830	Q49830 oryza sativ
93	39	51.3	304	2 Q7GCM2	Q7GCM2 oryza sativ
94	39	51.3	305	2 Q8Z299	Q8Z299 pyrobaculum
95	39	51.3	312	1 CCSA_ODOSI	CCSA_ODOSI odontella s
96	39	51.3	316	2 Q84QQ7	Q84QQ7 oryza sativ
97	39	51.3	318	2 Q9A6L6	Q9A6L6 caulobacter
98	39	51.3	327	2 Q93SG2	Q93SG2 eubacterium
99	39	51.3	335	2 Q9VTU8	Q9VTU8 drosophila
100	39	51.3	340	2 Q8FRSA	Q8FRSA corynebacte
101	39	51.3	411	2 Q7PDJ2	Q7PDJ2 plasmodium
102	39	51.3	437	2 Q9UFC5	Q9UFC5 homo sapten
103	39	51.3	492	2 Q6JAD9	Q6JAD9 zea mays (m
104	39	51.3	515	2 Q7PUS0	Q7PUS0 anopheles g

105	39	51.3	519	2	Q7UII0	Q7uii0 rhodopirell	178	38	50.0	564	2	Q6HFL4	Q6hfl4 bacillus th
106	39	51.3	547	2	Q7SEA8	Q7sea8 neurospora	179	38	50.0	566	2	Q7XQ98	Q7xq98 oryza sativ
107	39	51.3	560	2	P95227	P95227 mycobacteri	180	38	50.0	593	2	Q83CP3	Q83cp3 coxiella bu
108	39	51.3	560	2	Q92RX0	Q92rx0 rhizobium m	181	38	50.0	605	2	Q62OV3	Q62ov3 oryza sativ
109	39	51.3	560	2	Q7U2E6	Q7u2e6 mycobacteri	182	38	50.0	613	2	Q6LMB4	Q6lme4 photobacter
110	39	51.3	568	2	Q637N9	Q637n9 bacillus ce	183	38	50.0	627	2	Q49648	Q49648 arabidopsis
111	39	51.3	568	2	Q637P0	Q637p0 bacillus ce	184	38	50.0	628	2	Q74HV5	Q74hv5 lactobacill
112	39	51.3	568	2	Q733Q7	Q733q7 bacillus ce	185	38	50.0	647	2	Q8GSC2	Q8gac2 oryza sativ
113	39	51.3	568	2	Q81YA9	Q81ya9 bacillus an	186	38	50.0	657	2	Q8S5W4	Q8sw4 oryza sativ
114	39	51.3	568	2	Q81YB0	Q81yb0 bacillus an	187	38	50.0	688	2	Q8KBB9	Q8kbb9 chlorobium
115	39	51.3	568	2	Q6HFL0	Q6hfl0 bacillus th	188	38	50.0	699	2	Q8TWT4	Q8twt4 methanopyru
116	39	51.3	577	2	Q6X3X2	Q6x3x2 burkholderi	189	38	50.0	722	1	SWI2_SCHPO	Q10668 schizosacch
117	39	51.3	644	2	Q6DD72	Q6dd72 xenopus lae	190	38	50.0	723	2	Q9SV54	Q9sv54 arabidopsis
118	39	51.3	676	2	Q7RZM7	Q7rzmt7 neurospora	191	38	50.0	728	2	Q681A5	Q681a5 arabidopsis
119	39	51.3	741	2	Q6NNW6	Q6nnw6 drosophila	192	38	50.0	751	2	Q7S8X5	Q7s8x5 neurospora
120	39	51.3	750	2	Q8IH52	Q8ih52 drosophila	193	38	50.0	798	2	Q7FA54	Q7fa54 oryza sativ
121	39	51.3	865	2	Q7RIJ8	Q7rij8 giardia lam	194	38	50.0	831	2	Q9PU49	Q9pu49 gallus gall
122	39	51.3	887	2	Q9VSR1	Q9vsr1 drosophila	195	38	50.0	845	2	Q9LNQ0	Q9lnq0 arabidopsis
123	39	51.3	952	2	Q7XN31	Q7xn31 oryza sativ	196	38	50.0	901	2	Q18749	Q18749 caenorhabdi
124	39	51.3	1089	1	IMB3_YEAST	P32337 saccharomyc	197	38	50.0	931	2	Q6YS91	Q6ys91 oryza sativ
125	39	51.3	1290	2	Q82LI3	Q82li3 streptomyc	198	38	50.0	947	2	Q6SX96	Q6sx96 oryza sativ
126	39	51.3	1395	1	QGB3_HUMAN	Q8wml7 homo sapien	199	38	50.0	974	2	Q6PG27	Q6pg27 homo sapien
127	39	51.3	1467	2	Q9M033	Q9m033 arabidopsis	200	38	50.0	1504	2	Q68D66	Q68d66 homo sapien
128	39	51.3	8817	2	Q53840	Q53840 polyangium	201	38	50.0	1723	2	Q9Y4F4	Q9y4f4 homo sapien
129	38.5	50.7	239	2	Q67PV7	Q67pv7 symbiobacte	202	38	50.0	2493	2	Q6CMV4	Q6cmv4 kluyveromyc
130	38.5	50.7	327	2	Q82W5	Q82w5 nitrosomona	203	38	50.0	2518	2	Q8GBX7	Q8gbx7 polyangium
131	38	50.0	57	2	Q9KXK7	Q9kxk7 vibrio chol	204	38	50.0	2802	2	Q9DER5	Q9der5 gallus gall
132	38	50.0	71	1	RS21_PSEAE	Q9i5v8 pseudomonas	205	38	50.0	2870	2	Q9SQ02	Q9sq02 caenorhabdi
133	38	50.0	71	1	RS21_PSEPK	O51942 pseudomonas	206	38	50.0	3985	2	Q7Q293	Q7q293 anopheles g
134	38	50.0	71	1	RS21_PSESM	Q88a58 pseudomonas	207	37.5	49.3	368	2	Q8Y6F0	Q8y6f0 listeria mo
135	38	50.0	71	1	RS21_XANAC	P66535 xanthomonas	208	37.5	49.3	368	2	Q92AS1	Q92aa1 listeria in
136	38	50.0	71	1	RS21_XANCP	P66536 xanthomonas	209	37	48.7	74	2	Q9HRE2	Q9hre2 halobacteri
137	38	50.0	71	1	RS21_XYLFA	Q9p968 xylella fas	210	37	48.7	74	2	Q83BB9	Q83bb9 coxiella bu
138	38	50.0	71	1	RS21_XYLFT	Q87b16 xylella fas	211	37	48.7	79	2	Q8MIH0	Q8mih0 chaetosphae
139	38	50.0	73	2	Q7NUJ6	Q7nuj6 chromobacte	212	37	48.7	100	2	Q8X116	Q8x116 pseudomonas
140	38	50.0	74	2	Q6JIK6	Q6jik6 bacterioph	213	37	48.7	111	2	Q7QLD7	Q7qld7 anopheles g
141	38	50.0	74	2	Q77E99	Q77e99 bacterioph	214	37	48.7	121	2	Q7RZ9H	Q7rz9h neurospora
142	38	50.0	138	2	Q6W1Z2	Q6w1z2 rhizobium s	215	37	48.7	131	2	Q9A7L3	Q9a7l3 caulobacter
143	38	50.0	144	2	Q7MPE21	Q7mp21 vibrio vuln	216	37	48.7	136	2	Q6AITS	Q6ats5 oryza sativ
144	38	50.0	153	2	Q87IF3	Q87if3 vibrio para	217	37	48.7	150	2	Q7U775	Q7u775 synchococc
145	38	50.0	159	2	Q65OZ5	Q65oz5 oryza sativ	218	37	48.7	161	2	Q7NPD1	Q7npd1 gloeobacter
146	38	50.0	169	2	Q8Y8N1	Q8yn81 anabaena sp	219	37	48.7	186	2	Q6SM28	Q6sm28 bacillus li
147	38	50.0	175	2	Q61JN0	Q6ijn0 drosophila	220	37	48.7	187	1	LEPU_BACSU	P42959 bacillus su
148	38	50.0	180	2	Q7S6U1	Q7s6u1 neurospora	221	37	48.7	198	2	Q652K5	Q652k5 oryza sativ
149	38	50.0	186	2	Q8PXD9	Q8pxd9 methanosarc	222	37	48.7	202	2	Q8THA7	Q8tha7 methanosarc
150	38	50.0	190	2	Q87TJ5	Q87tj5 vibrio para	223	37	48.7	214	2	Q7ZFI1	Q7zfi1 desulfovibr
151	38	50.0	198	2	Q6L4M2	Q6l4m2 oryza sativ	224	37	48.7	219	2	Q7QDF7	Q7qdf7 anopheles g
152	38	50.0	217	2	Q8G837	Q8g837 bifidobacte	225	37	48.7	223	2	Q9AS87	Q9as87 oryza sativ
153	38	50.0	231	2	Q7U481	Q7u481 synchococc	226	37	48.7	224	2	Q8W845	Q8w845 todus angus
154	38	50.0	252	2	Q7N9C6	Q7n9c6 photorhabdu	227	37	48.7	224	2	Q8WC17	Q8wc17 todus angus
155	38	50.0	263	2	Q7NVN6	Q7nvnm6 chromobacte	228	37	48.7	229	2	Q8N579	Q8ns79 homo sapien
156	38	50.0	276	2	Q82UM4	Q82um4 nitrosomona	229	37	48.7	235	2	Q818H1	Q818h1 bacillus ce
157	38	50.0	278	2	Q6CRH9	Q6crh9 kluyveromyc	230	37	48.7	241	2	Q7RK90	Q7rk90 plasmodium
158	38	50.0	281	2	Q8S994	Q8s994 phaseolus a	231	37	48.7	249	2	Q6AG14	Q6agi4 leifsonia x
159	38	50.0	294	2	Q26941	Q26941 trypanosoma	232	37	48.7	256	2	Q9AF68	Q9af68 rhizobium l
160	38	50.0	307	2	Q9AT30	Q9at30 oryza sativ	233	37	48.7	265	2	Q67N88	Q67n88 symbiobacte
161	38	50.0	327	2	Q7P669	Q7p669 fusobacteri	234	37	48.7	268	2	Q6VAY4	Q6vay4 pseudomonas
162	38	50.0	327	2	Q8REA0	Q8rea0 fusobacteri	235	37	48.7	272	2	Q68ET6	Q68et6 xenopus lae
163	38	50.0	354	2	Q9RZU6	Q9rzue6 deinococcus	236	37	48.7	278	2	Q9H275	Q9h275 homo sapien
164	38	50.0	356	2	Q8H8C7	Q8hc7 oryza sativ	237	37	48.7	281	2	Q7XVX2	Q7xvx2 oryza sativ
165	38	50.0	357	2	Q75FZ4	Q75fz4 leptospira	238	37	48.7	281	2	Q7XW32	Q7xw32 oryza sativ
166	38	50.0	357	2	Q8EY06	Q8ey06 leptospira	239	37	48.7	281	2	Q88BI6	Q88bi6 pseudomonas
167	38	50.0	369	2	Q8RYR7	Q8ryr7 oryza sativ	240	37	48.7	286	1	PSS_MYCRO	P59949 mycobacteri
168	38	50.0	388	2	Q6NSF7	Q6nsf7 rhodopseudo	241	37	48.7	286	1	PSS_MYCTU	P96282 mycobacteri
169	38	50.0	403	2	Q7NFPQ1	Q7nfg1 gloeobacter	242	37	48.7	292	2	Q8WVL0	Q8wl10 homo sapien
170	38	50.0	411	2	Q9RU03	Q9ru03 deinococcus	243	37	48.7	297	2	Q71BE9	Q71be9 barley yell
171	38	50.0	515	2	Q94CG9	Q94cg9 zea mays (m	244	37	48.7	299	2	Q47236	Q47236 escherichia
172	38	50.0	516	2	Q9URR5	Q9urz5 penicillium	245	37	48.7	302	2	Q6SPR4	Q6spr4 oryza sativ
173	38	50.0	516	2	Q9URR6	Q9urr6 penicillium	246	37	48.7	310	1	DISP_MOUSE	Q9qy29 mus musculu
174	38	50.0	516	2	Q9URR7	Q9urr7 penicillium	247	37	48.7	311	2	Q8PD93	Q8pd93 xanthomonas
175	38	50.0	531	2	Q6K9W6	Q6k9w6 oryza sativ	248	37	48.7	323	2	Q94PC8	Q94pc8 coryphaena
176	38	50.0	548	2	Q9CD27	Q9cd27 mycobacteri	249	37	48.7	323	2	Q958V0	Q958v0 coryphaena
177	38	50.0	558	2	Q653S7	Q653s7 oryza sativ	250	37	48.7	323	2	Q958V1	Q958v1 coryphaena

251	37	48.7	323	2	Q958V2	Q958v2 coryphaena	324	37	48.7	1248	2	Q96762	O96762 trypanosoma
252	37	48.7	323	2	Q958V3	Q958v3 coryphaena	325	37	48.7	1333	2	Q95JT2	Q95jt2 arbidopsis
253	37	48.7	323	2	Q958V7	Q958v7 coryphaena	326	37	48.7	1356	2	Q95H77	Q95h77 arbidopsis
254	37	48.7	323	2	Q958X8	Q958x8 coryphaena	327	37	48.7	1430	2	Q9V4J4	Q9v4j4 drosophila
255	37	48.7	323	2	Q958X9	Q958x9 coryphaena	328	37	48.7	2412	1	POL1_BAYMG	Q04574 b genome po
256	37	48.7	323	2	Q958Y0	Q958y0 coryphaena	329	37	48.7	3149	1	TEGU_EBV	P03186 epstein-bar
257	37	48.7	323	2	Q958Y1	Q958y1 coryphaena	330	37	48.7	3149	2	Q777G4	Q777g4 human herpe
258	37	48.7	323	2	Q958Y2	Q958y2 coryphaena	331	37	48.7	3179	2	Q8V2A4	Q8v2a4 human herpe
259	37	48.7	325	2	Q8Z2L0	Q8zzl0 pyrobaculum	332	37	48.7	3195	2	Q6UDJ5	Q6udj5 psittacid h
260	37	48.7	330	2	Q7WYF3	Q7wyf3 pseudomonas	333	36.5	48.0	685	2	Q7ZWR9	Q7zwr9 xenopus lae
261	37	48.7	335	2	Q62A92	Q62a92 burkholderi	334	36.5	48.0	839	2	Q8LJZ0	Q8ljz0 sorghum bic
262	37	48.7	335	2	Q63NR6	Q63nr6 burkholderi	335	36	47.4	50	2	Q8Z1Y9	Q8z1y9 borrelia bu
263	37	48.7	344	2	Q73VDS	Q73vds mycobacteri	336	36	47.4	53	2	Q6IML8	Q6iml8 drosophila
264	37	48.7	353	2	Q62FG5	Q62fg5 burkholderi	337	36	47.4	54	2	Q8F2F6	Q8f2f6 oryza sativ
265	37	48.7	353	2	Q63VE1	Q63ve1 burkholderi	338	36	47.4	60	2	Q8XQA0	Q8xqa0 ralstonia s
266	37	48.7	357	2	Q8VB60	Q8vb60 white spot	339	36	47.4	69	2	Q83HD7	Q83hd7 tropheryma
267	37	48.7	357	2	Q91LR5	Q91lr5 white spot	340	36	47.4	88	2	Q938K9	Q938k9 temperate p
268	37	48.7	368	2	Q8P414	Q8p414 xanthomonas	341	36	47.4	88	2	Q79XU2	Q79xu2 streptococc
269	37	48.7	381	2	Q8ZXP7	Q8zxp7 pyrobaculum	342	36	47.4	88	2	Q8P0A0	Q8p0a0 streptococc
270	37	48.7	381	2	Q84ES2	Q84es2 wautersia o	343	36	47.4	88	2	Q7CN26	Q7cna26 streptococc
271	37	48.7	382	2	Q9NE87	Q9ne87 leishmania	344	36	47.4	88	2	Q6DEA0	Q6dea0 xenopus lae
272	37	48.7	395	2	Q8WV63	Q8wv63 homo sapien	345	36	47.4	96	2	Q9S552	Q9s552 borrelia bu
273	37	48.7	397	2	Q8P3V1	Q8p3v1 xanthomonas	346	36	47.4	100	2	Q8KH25	Q8kh25 pseudomonas
274	37	48.7	397	2	Q8PFD0	Q8pfd0 xanthomonas	347	36	47.4	101	2	Q8PG39	Q8pg39 xanthomonas
275	37	48.7	412	2	Q7CTH6	Q7cth6 agrobacteri	348	36	47.4	104	2	Q8YYM0	Q8yym0 anabaena sp
276	37	48.7	413	2	Q7NS03	Q7ns03 chromobacte	349	36	47.4	109	2	Q9Z387	Q9z387 borrelia bu
277	37	48.7	416	2	Q9RB05	Q9rb05 alcaligenes	350	36	47.4	111	2	Q8S077	Q8s077 oryza sativ
278	37	48.7	417	2	Q62173	Q62173 burkholderi	351	36	47.4	111	2	Q942U2	Q942u2 oryza sativ
279	37	48.7	417	2	Q63RG7	Q63rg7 burkholderi	352	36	47.4	124	2	Q8BP39	Q8bp39 mus musculu
280	37	48.7	418	2	Q67NS0	Q67ns0 symbiobacte	353	36	47.4	146	2	Q8EQ42	Q8eq42 oryza sativ
281	37	48.7	419	1	Y4WL_RHISN	P55571 rhizobium s	354	36	47.4	148	2	Q69QE8	Q69qe8 oryza sativ
282	37	48.7	420	2	Q92PJ9	Q92pj9 rhizobium m	355	36	47.4	150	2	Q9X917	Q9x917 streptomyce
283	37	48.7	422	2	Q8RZ59	Q8rz59 oryza sativ	356	36	47.4	155	2	Q7MK08	Q7mk08 vibrio vuln
284	37	48.7	440	2	Q91LG5	Q91lg5 white spot	357	36	47.4	155	2	Q8DA71	Q8da71 vibrio vuln
285	37	48.7	447	2	Q6CIJ8	Q6cij8 kluyveromyc	358	36	47.4	156	2	Q6AD72	Q6ad72 leifsonia x
286	37	48.7	458	1	YVTA_BACSU	Q8r911 bacillus su	359	36	47.4	163	2	Q8Z1Z2	Q8z1z2 borrelia bu
287	37	48.7	459	1	Y97E_TBSPA	Q83941 treponema p	360	36	47.4	165	2	Q8YXP3	Q8ypx3 anabaena sp
288	37	48.7	479	2	Q96SU8	Q96su8 homo sapien	361	36	47.4	172	2	Q6CGT6	Q6cgt6 yarrowia li
289	37	48.7	480	2	Q8RXTX2	Q8rtx2 uncultured	362	36	47.4	178	2	Q8WM50	Q8wm50 xenopus lae
290	37	48.7	483	2	Q9N0F4	Q9n0f4 mustela put	363	36	47.4	191	2	Q9ZNX5	Q9znx5 camellia si
291	37	48.7	504	2	Q6BXK0	Q6bxk0 debaryomyce	364	36	47.4	192	2	Q6CB84	Q6cb84 yarrowia li
292	37	48.7	507	2	Q9BVA8	Q9bva8 homo sapien	365	36	47.4	192	2	Q83FQ9	Q83fq9 tropheryma
293	37	48.7	508	2	Q9PFL57	Q9pl57 chlamydia m	366	36	47.4	193	2	Q7PQT2	Q7pqt2 anopheles g
294	37	48.7	513	2	Q90YAS	Q90yas anguilla ja	367	36	47.4	195	2	Q8NZ09	Q8nz09 streptococc
295	37	48.7	517	2	Q6RGPA4	Q6rgp4 streptomyce	368	36	47.4	195	2	Q99XJ8	Q99xj8 streptococc
296	37	48.7	528	2	Q92LY7	Q92ly7 helicobacte	369	36	47.4	195	2	Q8K511	Q8k511 streptococc
297	37	48.7	543	2	Q7K5M3	Q7k5m3 drosophila	370	36	47.4	198	2	Q6ANJ0	Q6anj0 desulfotale
298	37	48.7	565	2	Q73TQ3	Q73td3 mycobacteri	371	36	47.4	201	2	Q8TLV5	Q8tlv5 methanosarc
299	37	48.7	566	2	Q69TK3	Q69tk3 oryza sativ	372	36	47.4	204	2	Q6D313	Q6d313 erwinita car
300	37	48.7	576	2	Q6DCW8	Q6dcw8 xenopus lae	373	36	47.4	215	2	Q7XY83	Q7xy83 griffithsia
301	37	48.7	596	2	Q6F2Y6	Q6f2y6 oryza sativ	374	36	47.4	217	2	Q8PIZ1	Q8piz1 xanthomonas
302	37	48.7	601	2	Q9X8H7	Q9x8h7 streptomyce	375	36	47.4	222	2	Q655P4	Q655p4 oryza sativ
303	37	48.7	607	2	Q25379	Q25379 helicobacte	376	36	47.4	224	2	Q8DPY9	Q8dpy9 vibrio vuln
304	37	48.7	610	2	Q7VEF7	Q7vef7 prochloroco	377	36	47.4	233	2	Q7XTV6	Q7xtv6 oryza sativ
305	37	48.7	613	2	Q8QTE5	Q8qte5 white spot	378	36	47.4	235	2	Q9XTH1	Q9xth1 caenorhabdi
306	37	48.7	636	2	Q8PC9A9	Q8pca9 xanthomonas	379	36	47.4	238	2	Q6NFB6	Q6nfb6 corynebacte
307	37	48.7	636	2	Q8PNZ1	Q8pnz1 xanthomonas	380	36	47.4	246	2	Q74606	Q746q6 geobacter s
308	37	48.7	668	2	Q7KR89	Q7kr89 drosophila	381	36	47.4	248	2	Q8PBN3	Q8pbn3 xanthomonas
309	37	48.7	675	2	Q6C4A2	Q6c4a2 yarrowia li	382	36	47.4	249	1	Y8G0_STRO	P40179 streptomyce
310	37	48.7	719	2	Q6C496	Q6c496 yarrowia li	383	36	47.4	249	2	Q85AG7	Q85ag7 yersinia pe
311	37	48.7	763	2	Q941Z6	Q941z6 oryza sativ	384	36	47.4	249	2	Q65KV4	Q65kv4 bacillus li
312	37	48.7	769	1	DCB2_MOUSE	Q91zv3 mus musculu	385	36	47.4	251	2	Q9D9K5	Q9d9k5 m mus muscu
313	37	48.7	769	1	DCB2_RAT	Q91zv2 rattus norv	386	36	47.4	252	2	Q9BQ73	Q9bq73 homo sapien
314	37	48.7	772	2	Q8SPR7	Q8spr7 sus scrofa	387	36	47.4	253	2	Q73HU0	Q73hu0 wolbachia p
315	37	48.7	775	1	DCB2_HUMAN	Q96pd2 homo sapien	388	36	47.4	256	2	Q8A2S0	Q8a2s0 caulobacter
316	37	48.7	821	2	Q6LNE0	Q6lne0 photobacter	389	36	47.4	257	2	Q6YWL4	Q6ywl4 oryza sativ
317	37	48.7	835	1	TSD2_USTWA	Q99107 ustilago ma	390	36	47.4	259	1	YGFH_ECOLI	P64558 escherichia
318	37	48.7	939	2	Q49467	Q49467 arbidopsis	391	36	47.4	259	1	YGFH_ECOLI	P64557 escherichia
319	37	48.7	1007	2	Q9F3J0	Q9f3j0 streptomyce	392	36	47.4	259	2	Q8VK87	Q8vk87 anabaena sp
320	37	48.7	1038	2	Q8XR28	Q8xr28 ralstonia s	393	36	47.4	259	2	Q8FE83	Q8fe83 escherichia
321	37	48.7	1075	2	Q73RD3	Q73rd3 treponema d	394	36	47.4	264	2	Q71LV2	Q71lv2 melospiza m
322	37	48.7	1148	2	Q675R7	Q675r7 oikopleura	395	36	47.4	264	2	Q9LG70	Q9lg70 oryza sativ
323	37	48.7	1189	2	Q85152	Q85152 photorhabdu	396	36	47.4	268	2	Q8QM77	Q8qmy7 cowpox viru

397	36	47.4	269	1	VI03_VACCV	P12923 vaccinia vi	470	36	47.4	458	2	O7XWC1	O7xwc1 oryza sativ
398	36	47.4	269	1	VI03_VARV	P33000 variola vir	471	36	47.4	462	2	O8TPE0	O8tpe0 methanosarc
399	36	47.4	269	2	O8V516	O8v516 monkeypox v	472	36	47.4	466	2	O7VVF3	O7vvf3 bordetella
400	36	47.4	269	2	O6RZN0	O6rzn0 rabbitpox v	473	36	47.4	468	2	O89XMS	O89xms bradyrhizob
401	36	47.4	269	2	O76Q30	O76q30 variola min	474	36	47.4	475	2	O82BQ8	O82bq8 streptomyce
402	36	47.4	269	2	O76QK8	O76qk8 cowpox viru	475	36	47.4	476	2	O19657	O19657 caenorhabdi
403	36	47.4	269	2	O76ZV0	O76zv0 vaccinia vi	476	36	47.4	478	1	5HT3_HUMAN	P46098 homo sapien
404	36	47.4	270	2	O93UX4	O93ux4 agrobacteri	477	36	47.4	480	2	O82FI6	O82fi6 streptomyce
405	36	47.4	271	2	O8VZU8	O8vzu8 camelpox vi	478	36	47.4	482	2	O6NWI1	O6nwi1 homo sapien
406	36	47.4	271	2	O775X1	O775x1 camelpox vi	479	36	47.4	484	2	O7KZM7	O7kzm7 homo sapien
407	36	47.4	273	1	MURI_AGRRT5	O8ue93 agrobacteri	480	36	47.4	485	1	GLGA_BACST	O08328 bacillus st
408	36	47.4	274	2	O49G1	O49g1 oryza sativ	481	36	47.4	487	2	O46080	O46080 drosophila
409	36	47.4	274	2	O67LR5	O67lr5 symbiobacte	482	36	47.4	488	2	O7WB87	O7wb87 bordetella
410	36	47.4	276	2	O8JTV6	O8jtv6 lumpy skin	483	36	47.4	488	2	O7WMQ5	O7wmq5 bordetella
411	36	47.4	285	2	O91MW2	O91mw2 lumpy skin	484	36	47.4	490	2	O7EYM6	O7eym6 oryza sativ
412	36	47.4	285	2	O8FLW9	O8flw9 escherichia	485	36	47.4	493	2	O8W0R0	O8w0r0 sorghum bic
413	36	47.4	286	2	O6ARK3	O6ark3 desulfotale	486	36	47.4	497	2	O17513	O17513 caenorhabdi
414	36	47.4	287	2	O65RQ0	O65rq0 manheimia	487	36	47.4	502	2	O8FBL4	O8fbl4 escherichia
415	36	47.4	293	2	O6LFX7	O6lfx7 photobacter	488	36	47.4	508	1	YM05_ARCFU	O28078 archaeoglob
416	36	47.4	294	2	O67L16	O67l16 symbiobacte	489	36	47.4	509	2	O52375	O52375 pseudomonas
417	36	47.4	297	2	O9ACY0	O9acy0 streptomyce	490	36	47.4	509	2	O87W76	O87w76 pseudomonas
418	36	47.4	298	2	O8UNG7	O8uwn7 gallus gall	491	36	47.4	513	2	O7MG66	O7mg66 vibrio vuln
419	36	47.4	320	2	O7N6G1	O7n6g1 photorhabdu	492	36	47.4	526	2	O6D271	O6d271 erwinia car
420	36	47.4	300	2	O8NKT0	O8nkt0 thermococcu	493	36	47.4	530	2	O69IM5	O69im5 oryza sativ
421	36	47.4	300	2	O9HHC6	O9hnc6 thermococcu	494	36	47.4	530	2	O88W48	O88w48 lactobacill
422	36	47.4	304	2	O96182	O96i82 homo sapien	495	36	47.4	531	2	O7VVV2	O7vvv2 bordetella
423	36	47.4	304	2	O6ZMB1	O6zmb1 homo sapien	496	36	47.4	538	2	O9KY53	O9ky53 streptomyce
424	36	47.4	305	2	O9KFT1	O9kft1 bacillus ha	497	36	47.4	565	2	O8L8C4	O8l8c4 arabisdopsis
425	36	47.4	320	2	O8BP92	O8bp92 mus muscullu	498	36	47.4	565	2	O6DR03	O6dr03 arabisdopsis
426	36	47.4	321	2	O70341	O70341 mus muscullu	499	36	47.4	567	2	O22814	O22814 arabisdopsis
427	36	47.4	327	2	O6IND9	O6ind9 xenopus lae	500	36	47.4	567	2	O81AI2	O81ai2 bacillus ce
428	36	47.4	329	2	O9V2X6	O9v2x6 methanobact	501	36	47.4	576	1	PYRG_ASHGO	O75117 ashbya goss
429	36	47.4	330	2	O8H2P0	O8h2p0 oryza sativ	502	36	47.4	578	2	O8BPX9	O8bpx9 mus muscullu
430	36	47.4	330	2	O842G5	O842g5 pseudomonas	503	36	47.4	578	2	O9WU80	O9wu80 mus muscullu
431	36	47.4	331	2	O8TVS6	O8tvs6 methanopyru	504	36	47.4	582	2	O924V4	O924v4 rattus norv
432	36	47.4	331	2	O9RRL2	O9rrl2 deinococcus	505	36	47.4	592	1	STB3_HUMAN	O00186 homo sapien
433	36	47.4	335	2	O6L4Y1	O6l4y1 oryza sativ	506	36	47.4	595	2	O7S9F6	O7s9f6 neurospora
434	36	47.4	335	2	O9F267	O9f267 actinobacil	507	36	47.4	595	2	O9SDM4	O9sdm4 dunaliella
435	36	47.4	336	2	O89W09	O89w09 bradyrhizob	508	36	47.4	599	2	O67RH1	O67rh1 symbiobacte
436	36	47.4	340	2	O622B3	O622b3 oryza sativ	509	36	47.4	601	2	O82EE2	O82ee2 streptomyce
437	36	47.4	347	2	O91E14	O91e14 rice black	510	36	47.4	605	2	O7XIN9	O7xin9 oryza sativ
438	36	47.4	352	2	O9BXA0	O9bxa0 homo sapien	511	36	47.4	613	2	O9FTG6	O9ftg6 oryza sativ
439	36	47.4	352	2	O66977	O66977 versinia ps	512	36	47.4	628	2	O6FVA6	O6fva6 candida gla
440	36	47.4	352	2	O8ZDV5	O8zdv5 versinia pe	513	36	47.4	639	2	O7QUK1	O7quk1 giardia lam
441	36	47.4	355	2	O9LZU3	O9lzu3 arabisdopsis	514	36	47.4	635	2	O9CH87	O9ch87 lactococcus
442	36	47.4	355	2	O92X89	O92xe9 rhizobium m	515	36	47.4	697	2	O9CA19	O9ca19 arabisdopsis
443	36	47.4	358	1	AROG_BACSU	P39912 b aroa(g) p	516	36	47.4	714	2	O9VQV3	O9vqv3 drosophila
444	36	47.4	361	2	O8D0R2	O8d0r2 versinia pe	517	36	47.4	718	1	PERM_MOUSE	P11247 mus muscullu
445	36	47.4	371	2	O9U3P5	O9u3p5 caenorhabdi	518	36	47.4	730	2	O06634	O06634 bovine herp
446	36	47.4	373	2	O9XXK7	O9xxk7 caenorhabdi	519	36	47.4	735	1	AMDR_ASPOR	O06157 aspergillus
447	36	47.4	378	2	O7MJ80	O7mj80 vibrio vuln	520	36	47.4	769	1	COMP_BACSU	O09027 bacillus su
448	36	47.4	378	2	O8DAN0	O8dan0 vibrio vuln	521	36	47.4	819	2	O88G18	O88g18 pseudomonas
449	36	47.4	389	2	O8PFJ6	O8pfj6 xanthomonas	522	36	47.4	847	2	O8YOC4	O8yoc4 ralatonia s
450	36	47.4	391	2	O8RWI0	O8rwi0 arabisdopsis	523	36	47.4	880	2	O8GJN2	O8gjn2 streptomyce
451	36	47.4	391	2	O9PPD7	O9ppd7 arabisdopsis	524	36	47.4	939	1	XPC_HUMAN	O01831 homo sapien
452	36	47.4	395	2	O6ND83	O6nd83 rhodopseudo	525	36	47.4	940	2	O96AX0	O96ax0 homo sapien
453	36	47.4	410	2	O71098	O71098 bovine aden	526	36	47.4	1001	2	O8N283	O8n283 homo sapien
454	36	47.4	420	2	O7NSV5	O7nav5 chromobacte	527	36	47.4	1021	2	O6GTQ7	O6gtq7 mus muscullu
455	36	47.4	420	2	O6AN13	O6an13 desulfotale	528	36	47.4	1029	2	O6P5F1	O6p5f1 mus muscullu
456	36	47.4	421	2	O82019	O82ql9 streptomyce	529	36	47.4	1055	1	EPB2_HUMAN	P29323 homo sapien
457	36	47.4	422	2	O9RJV8	O9rvj8 streptomyce	530	36	47.4	1070	1	RPOB_CHAGL	O8ma12 chaetophae
458	36	47.4	424	2	O6A201	O6a201 oenococcus	531	36	47.4	1080	2	O9SDC1	O9sdc1 oryza sativ
459	36	47.4	431	2	O8U1I9	O8u1i9 agrobacteri	532	36	47.4	1098	2	O9RQ31	O9rq31 clostridium
460	36	47.4	432	2	O751K5	O751k5 ashbya goss	533	36	47.4	1158	2	O7Q7D6	O7q7d6 anopheles g
461	36	47.4	434	1	YK12_YEAST	P36080 saccharomyc	534	36	47.4	1175	2	O9X716	O9x716 clostridium
462	36	47.4	435	2	O7D1D9	O7d1d9 agrobacteri	535	36	47.4	1205	2	O6ZU48	O6zu48 homo sapien
463	36	47.4	438	2	O6N9L1	O6n9l1 rhodopseudo	536	36	47.4	1281	2	O01789	O01789 caenorhabdi
464	36	47.4	440	2	O88VG0	O88vg0 lactobacill	537	36	47.4	1282	2	O7RY84	O7ry84 neurospora
465	36	47.4	442	2	O9C1K8	O9c1k8 neurospora	538	36	47.4	1326	2	O23829	O23829 calliphora
466	36	47.4	445	1	ARLY_XYLFA	O9pem5 xylella fas	539	36	47.4	1343	2	O06635	O06635 bovine herp
467	36	47.4	447	1	O655P0	O655p0 oryza sativ	540	36	47.4	1353	1	XDH_CALVI	P08793 calliphora
468	36	47.4	451	1	CBPS_STRGR	P18143 streptomyce	541	36	47.4	1363	2	O9M1F5	O9m1f5 arabisdopsis
469	36	47.4	452	2	O8L277	O8l277 proteus vul	542	36	47.4	1501	2	O9SD86	O9sd86 arabisdopsis



543	36	47.4	1706	2	O42900	O42900 schizosacch	616	35	46.1	215	2	Q8EQA1	Q8eqa1 oceanobacil
544	36	47.4	1722	1	RBE2_HUMAN	P29375 homo sapien	617	35	46.1	219	2	Q89DR6	Q89dr6 bradyrhizob
545	36	47.4	1842	2	Q9EWP4	Q9ewp4 streptomyc	618	35	46.1	220	2	Q6IK12	Q6ik12 drosophila
546	36	47.4	1883	2	Q8LMR2	Q8lmr2 oryza sativ	619	35	46.1	224	2	Q69836	Q69836 streptomyc
547	36	47.4	1909	1	DICE ARATH	Q9sp32 arabidopsis	620	35	46.1	227	2	Q7J3Z2	Q7j3z2 scomberom
548	36	47.4	4355	2	Q9P2D7	Q9p2d7 homo sapien	621	35	46.1	227	2	Q7J3Z8	Q7j3z8 scomberom
549	35.5	46.7	189	2	Q96J24	Q96j24 rhizobium l	622	35	46.1	227	2	Q7J405	Q7j405 scomberom
550	35.5	46.7	219	2	P96667	P96667 bacillus su	623	35	46.1	227	2	Q7J409	Q7j409 scomberom
551	35.5	46.7	282	2	Q7QRC6	Q7qrc6 giardia lam	624	35	46.1	227	2	Q9TGC4	Q9tgc4 scomberom
552	35.5	46.7	360	2	Q9F889	Q9f889 salmonella	625	35	46.1	227	2	Q9TGC5	Q9tgc5 scomberom
553	35.5	46.7	363	1	OMP_SALTI	Q56113 salmonella	626	35	46.1	227	2	Q9TGC6	Q9tgc6 scomberom
554	35.5	46.7	363	1	OMP_SALTY	P37432 salmonella	627	35	46.1	227	2	Q9TGC7	Q9tgc7 scomberom
555	35.5	46.7	363	2	Q9K3E7	Q9k3e7 salmonella	628	35	46.1	227	2	Q9TGC8	Q9tgc8 scomberom
556	35.5	46.7	3105	2	Q8UZI9	Q8uzi9 cercaripthec	629	35	46.1	227	2	Q9TGC8	Q9tgc8 scomberom
557	35	46.1	43	1	PSBY_SYNEL	Q8dkm3 synecchococ	630	35	46.1	227	2	Q9XJZ4	Q9xjz4 scomberom
558	35	46.1	63	2	Q7SM72	Q7sm72 oryza sativ	631	35	46.1	227	2	Q9XJZ6	Q9xjz6 scomberom
559	35	46.1	74	2	Q8P1Q2	Q8plq2 streptococ	632	35	46.1	227	2	Q9XK27	Q9xk27 scomberom
560	35	46.1	84	2	Q6EQX1	Q6eqx1 oryza sativ	633	35	46.1	227	2	Q9XPA4	Q9xpa4 scomberom
561	35	46.1	86	2	Q6UQ21	Q6uq21 uncultured	634	35	46.1	227	2	Q9XPA9	Q9xpa9 scomberom
562	35	46.1	86	2	Q8VB28	Q8vb28 white spot	635	35	46.1	227	2	Q9XPB0	Q9xpb0 scomberom
563	35	46.1	88	2	Q8K6A1	Q8k6a1 streptococ	636	35	46.1	227	2	Q9XPB1	Q9xpb1 scomberom
564	35	46.1	88	2	Q9A0P0	Q9a0p0 streptococ	637	35	46.1	227	2	Q8GRV7	Q8grv7 oryza sativ
565	35	46.1	88	2	Q7CWM8	Q7cmw8 streptococ	638	35	46.1	227	2	Q6PWQ4	Q6pwq4 xenopus lae
566	35	46.1	91	2	Q31460	Q31460 bacillus su	639	35	46.1	228	2	Q64Z69	Q64z69 bacteroides
567	35	46.1	91	2	Q833C0	Q833c0 enterococcu	640	35	46.1	229	2	Q8A1L4	Q8a1l4 bacteroides
568	35	46.1	93	2	Q620I2	Q620i2 oryza sativ	641	35	46.1	236	2	Q8NMB1	Q8nmb1 corynebacte
569	35	46.1	93	2	Q8KFG0	Q8kfg0 chlorobium	642	35	46.1	237	2	Q83A55	Q83a55 coxiellia bu
570	35	46.1	109	2	Q6K385	Q6k385 oryza sativ	643	35	46.1	238	2	Q8FMB8	Q8fmb8 corynebacte
571	35	46.1	110	2	Q6TVT2	Q6tv22 orf virus.	644	35	46.1	239	2	Q8FS72	Q8fsh2 corynebacte
572	35	46.1	110	2	Q6TW62	Q6tw62 orf virus.	645	35	46.1	240	2	Q7JWM0	Q7jwm0 mycobacteri
573	35	46.1	117	2	Q94DK6	Q94dk6 oryza sativ	646	35	46.1	243	2	Q8EV53	Q8ev53 mycoplasma
574	35	46.1	120	2	Q8DDT5	Q8ddt5 vibrio vuln	647	35	46.1	245	2	Q8WM51	Q8wm51 xenopus lae
575	35	46.1	129	2	Q8H7Y2	Q8h7y2 oryza sativ	648	35	46.1	246	2	Q67TN7	Q67ln7 oryza sativ
576	35	46.1	131	2	Q8V748	Q8v748 listeria mo	649	35	46.1	247	2	Q8BXP1	Q8bxp1 mus musculu
577	35	46.1	131	2	Q92BP6	Q92bp6 listeria in	650	35	46.1	248	1	TRPC_SULSO	Q0i121 sulfolobus
578	35	46.1	131	2	Q71ZK6	Q71zk6 listeria mo	651	35	46.1	248	2	Q6NJS9	Q6nj9 corynebacte
579	35	46.1	140	2	Q8WV74	Q8wv74 homo sapien	652	35	46.1	251	2	Q6N4A9	Q6n4a9 rhodopseuo
580	35	46.1	140	2	Q6K549	Q6k549 oryza sativ	653	35	46.1	252	2	Q7NTD4	Q7ntd4 chromobacte
581	35	46.1	142	2	Q8BV45	Q8bv45 mus musculu	654	35	46.1	253	2	Q9CWB5	Q9cwe5 mus musculu
582	35	46.1	143	2	Q49113	Q49113 arabidopsis	655	35	46.1	253	2	Q8BXV2	Q8bxv2 mus musculu
583	35	46.1	143	2	Q8BWA7	Q8bwa7 arabidopsis	656	35	46.1	254	2	Q8U9H7	Q8u9h7 agrobacteri
584	35	46.1	145	2	Q6LSD5	Q6led5 photobacter	657	35	46.1	255	2	Q9LIY6	Q9liy6 oryza sativ
585	35	46.1	150	2	Q8TS13	Q8ts13 methanosarc	658	35	46.1	255	2	Q7NJM2	Q7njm2 gloeobacter
586	35	46.1	151	2	Q686C4	Q686c4 pseudomonas	659	35	46.1	257	1	HB2P_RABIT	P20756 oryctolagus
587	35	46.1	151	2	Q686F1	Q686f1 pseudomonas	660	35	46.1	257	2	Q82M08	Q82mq8 streptomyc
588	35	46.1	155	2	Q8GYM7	Q8gym7 arabidopsis	661	35	46.1	261	2	Q7D546	Q7d546 mycobacteri
589	35	46.1	155	2	Q34694	Q34694 bacillus su	662	35	46.1	262	2	Q8GAG9	Q8gag9 arthrobacte
590	35	46.1	157	2	Q8S778	Q8s778 oryza sativ	663	35	46.1	264	2	Q752V9	Q752v9 ashbya goss
591	35	46.1	157	2	Q7XFS2	Q7xf22 oryza sativ	664	35	46.1	264	2	Q655W1	Q655w1 oryza sativ
592	35	46.1	168	1	SULA_SERMA	P08845 serratia ma	665	35	46.1	264	2	Q6DH50	Q6dh50 brachydanio
593	35	46.1	168	2	Q9UX13	Q9ux13 sulfolobus	666	35	46.1	269	1	V103_VACCC	P20499 vaccinia vi
594	35	46.1	170	2	Q7V826	Q7v826 prochloroco	667	35	46.1	269	2	Q9D7E2	Q9d7e2 mus musculu
595	35	46.1	172	1	TRPP_BACSU	O07515 bacillus su	668	35	46.1	269	2	Q9QYL7	Q9qyl7 mus musculu
596	35	46.1	173	2	Q8TWZ8	Q8twz8 methanopyru	669	35	46.1	269	2	Q931L8	Q931l8 vaccinia vi
597	35	46.1	183	2	Q6ETW4	Q6etw4 oryza sativ	670	35	46.1	269	2	Q8JLF3	Q8jlf3 actinomelia
598	35	46.1	184	2	Q8LGB8	Q8lgb8 arabidopsis	671	35	46.1	269	2	Q9JFC7	Q9jfc7 vaccinia vi
599	35	46.1	184	2	Q9CAF9	Q9caf9 arabidopsis	672	35	46.1	271	1	CDSA_PSEAE	Q59640 p phosphati
600	35	46.1	190	2	Q9NLA0	Q9nla0 podocoryne	673	35	46.1	271	2	Q88MH5	Q88mh5 pseudomonas
601	35	46.1	192	2	Q8A700	Q8a700 bacteroides	674	35	46.1	272	2	Q79VT2	Q79vt2 salmonella
602	35	46.1	192	2	Q8A769	Q8a769 bacteroides	675	35	46.1	272	2	Q9R2H4	Q9r2h4 plasmodi r64
603	35	46.1	192	2	Q8JW05	Q8jw05 spring beau	676	35	46.1	274	2	Q9VNM8	Q9vnm8 drosophila
604	35	46.1	195	2	Q9FC16	Q9fc16 streptomyc	677	35	46.1	277	2	Q7S8H9	Q7s8h9 neurospora
605	35	46.1	198	2	Q89HG8	Q89hg8 bradyrhizob	678	35	46.1	284	2	Q8EV55	Q8ev55 mycoplasma
606	35	46.1	202	2	Q8TDY7	Q8tdy7 homo sapien	679	35	46.1	286	2	Q6ZDW3	Q6zdw3 oryza sativ
607	35	46.1	203	2	Q6ESG2	Q6esg2 oryza sativ	680	35	46.1	286	2	Q6S888	Q6s888 barley yell
608	35	46.1	203	2	Q6N204	Q6n204 rhodopseuo	681	35	46.1	286	2	Q6S889	Q6s889 barley yell
609	35	46.1	206	2	Q8WUY8	Q8wuy8 homo sapien	682	35	46.1	287	2	Q6S890	Q6s890 barley yell
610	35	46.1	206	2	Q9NS72	Q9ns72 homo sapien	683	35	46.1	287	1	RL6_HUMAN	Q02878 homo sapien
611	35	46.1	206	2	Q9N0D0	Q9n0d0 macaca faec	684	35	46.1	288	2	Q8N5Z7	Q8n5z7 homo sapien
612	35	46.1	207	2	Q8BVG8	Q8bv8 m mus muscu	685	35	46.1	288	2	Q8TBK5	Q8tbk5 homo sapien
613	35	46.1	207	2	Q82H72	Q82h72 streptomyc	686	35	46.1	288	2	Q8WW97	Q8ww97 homo sapien
614	35	46.1	210	2	Q9VY84	Q9vy84 drosophila	687	35	46.1	288	2	Q96463	Q96463 hordeum vul
615	35	46.1	211	2	Q7Q127	Q7q127 anopheles g	688	35	46.1	288	2	Q6QM24	Q6qm24 chinchilla

```

689 35 46.1 289 2 Q9HBB3
690 35 46.1 290 2 Q815C1
691 35 46.1 290 2 Q8NTY6
692 35 46.1 292 2 Q6M8J2
693 35 46.1 293 2 Q7Q8S8
694 35 46.1 294 1 DAPA_RALSO
695 35 46.1 294 2 Q9FK20
696 35 46.1 297 2 Q826C1
697 35 46.1 307 2 Q67RE1
698 35 46.1 307 2 Q7MKK1
699 35 46.1 308 2 Q46370
700 35 46.1 310 1 DHYS_THEAC
701 35 46.1 310 2 Q6A7F8
702 35 46.1 310 2 Q9RIV5
703 35 46.1 312 2 Q8S5E5
704 35 46.1 312 2 Q7XG48
705 35 46.1 315 2 Q7B3G3
706 35 46.1 315 2 Q73VJ0
707 35 46.1 316 2 Q7CXN8
708 35 46.1 316 2 Q8BY54
709 35 46.1 317 2 P95110
710 35 46.1 317 2 Q7TXH7
711 35 46.1 321 2 Q815P3
712 35 46.1 323 2 Q6XYZ2
713 35 46.1 324 2 Q8UDE2
714 35 46.1 325 2 Q8UHf8
715 35 46.1 325 2 Q89YK8
716 35 46.1 327 1 KDG7_ECO57
717 35 46.1 327 1 KDG7_ECOLI
718 35 46.1 333 2 Q7D0U2
719 35 46.1 335 2 Q97DC5
720 35 46.1 339 1 KDG7_ERWCH
721 35 46.1 345 2 Q6CK03
722 35 46.1 345 2 Q733Q3
723 35 46.1 348 2 Q8HMN3
724 35 46.1 348 2 Q8DH32
725 35 46.1 348 2 Q9A4G3
726 35 46.1 350 2 Q69628
727 35 46.1 350 2 Q7TW03
728 35 46.1 352 2 Q9BDS5
729 35 46.1 352 2 Q47071
730 35 46.1 356 2 Q65861
731 35 46.1 357 1 MURG_BORBR
732 35 46.1 357 1 MURG_BORPA
733 35 46.1 357 1 MURG_BORPE
734 35 46.1 364 2 Q700C7
735 35 46.1 374 2 P73953
736 35 46.1 381 2 Q7XW09
737 35 46.1 389 2 Q85150
738 35 46.1 390 1 PYRB_ARATH
739 35 46.1 390 2 Q94A47
740 35 46.1 402 2 Q57029
741 35 46.1 406 2 Q7VV29
742 35 46.1 407 2 Q9RI96
743 35 46.1 408 2 Q7V1B7
744 35 46.1 410 2 P72472
745 35 46.1 410 2 Q7WN77
746 35 46.1 410 2 Q9RJJ8
747 35 46.1 411 2 Q53873
748 35 46.1 414 2 Q8I6J5
749 35 46.1 417 2 Q7MQS8
750 35 46.1 423 2 Q8UC88

```

## ALIGNMENTS

```

RESULT 1
Q6CK41 PRELIMINARY; PRT; 420 AA.
ID AC Q6CK41
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)

```

```

DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Kluyveromyces lactis strain NRRL Y-1140 chromosome F of strain NRRL Y-
DE 1140 of Kluyveromyces lactis.
DE ORFNames=KLLA0F13728G;
OS Kluyveromyces lactis NRRL Y-1140.
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Kluyveromyces.
OX NCBI_TaxID=284590;
[1]
SEQUENCE FROM N.A.
RC STRAIN=NRRL Y-1140;
RG Genolevures;
RA Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S.,
RA Lafontaine I., de Montigny J., Marck C., Neuveglise C., Talla E., V.,
RA Goffard N., Frangoul L., Algie M., Anchoard V., Babour A., Barbe V.,
RA Barnay S., Blanchin S., Beckerich J.M., Beyne E., Bleykasten C.,
RA Boisrame A., Boyer E., Cattolico L., Confanioleri F., de Daruvar A.,
RA Despons L., Fabre J., Fairhead C., Ferry-Dumazet H., Groppi A.,
RA Hantraye F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,
RA Kerrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,
RA Nicaut J.M., Nikolski M., Ozias S., Ozier-Kalogeropoulos O.,
RA Pellenz S., Potier S., Richard G.F., Straub M.L., Suleau A.,
RA Swennene D., Tekala F., Wesolowski-Louvel M., Westhof E., Wirth B.,
RA Zenlou-Meyer M., Zivanovic I., Bolotin-Fukuhara M., Thierry A.,
RA Bouchier C., Caudron B., Scarpelli C., Gaillardin C., Weissenbach J.,
RA Wincker P., Souciet J.L.;
RT "Genome evolution in yeasts.";
RL Nature 430:35-44(2004).
[2]
SEQUENCE FROM N.A.
RC STRAIN=NRRL Y-1140;
RA Genoscope;
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; CR382126; CAG98406.1; -.
SQ SEQUENCE 420 AA; 47286 MW; 56B7A79313B4B452 CRC64;

Query Match 59.2%; Score 45; DB 2; Length 420;
Best Local Similarity 52.9%; Pred. No. 47;
Matches 9; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 1 KKKRKAANAALVLPVLLA 17
: |||||:: :|||
Db 66 RSKRKAASSTPPVVLLA 82

RESULT 2
Q7SBW3 PRELIMINARY; PRT; 157 AA.
ID Q7SBW3
AC Q7SBW3;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Predicted protein.
GN Name=NCU07845.1;
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
[1]
SEQUENCE FROM N.A.
RC STRAIN=OR74N;
RA Galagan J.E., Calvo S.E., Borkovich K.A., Selker E.U., Read N.D.,
RA Jaffe D., FitzHugh W., Ma L.-J., Smirnov S., Purcell S., Rehman B.,
RA Elkins T., Engels R., Wang S., Nielsen C.B., Butler J., Endrizzi M.,
RA Qui D., Ianakiev P., Pedersen D., Nelson M., Washburne M.,
RA Selitrennikoff C.P., Kinsey J.A., Braun E.L., Zelter A., Schulte U.,
RA Kothe G.O., Jedd G., Mewes W., Staben C., Marcotte E., Greenberg D.,
RA Roy A., Foley K., Naylor J., Thomann N., Barrett R., Gnerre S.,
RA Kamal M., Kamysyssel M., Mauceli E., Bielke C., Rudd S., Frisman D.,
RA Krystofova S., Rasmussen C., Metzberg R.L., Perkins D.D., Kroken S.,
RA Cogoni C., Macino G., Catchside D., Li W., Pratt R.J., Omani S.A.,
RA DeSouza C.C., Glass L., Orbach M.J., Berglund J., Voelker R.,
RA Yarden O., Plamann M., Seiler S., Dunlap J., Radford A., Aramayo R.,

```

```

RA Natvig D.O., Alex L.A., Mannheim G., Ebole D.J., Freitag M.,
RA Paulsen I., Sachs M.S., Lander E.S., Nussbaum C., Birren B.;
RA "The Genome Sequence of the Filamentous Fungus Neurospora crassa.";
RL Nature 0:0-0(2003).
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AABX01000162; EAA33870.1; -.
SQ SEQUENCE 157 AA; 17703 MW; C62F0E2DD3834F3B CRC64;

Query Match 57.9%; Score 44; DB 2; Length 157;
Best Local Similarity 64.3%; Pred. No. 28;
Matches 9; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 KKKRKAANAALLPV 14
Db 126 RRSRAAPALLPV 139

RESULT 3
QBLR32 PRELIMINARY; PRT; 242 AA.
AC QBLR32;
DT 01-OCT-2002 (TREMELrel. 22, Created)
DT 01-OCT-2002 (TREMELrel. 22, Last sequence update)
DT 01-MAR-2004 (TREMELrel. 26, Last annotation update)
DE P0671D01.19 protein.
GN Name=P0671D01.19;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=12447438; DOI=10.1039/nature01184;
RA Sasaki T., Matsumoto T., Yamamoto K., Sakata K., Baba T., Katayose Y.,
RA Wu J., Nimura Y., Cheng Z., Nagamura Y., Antonio B.A., Kanamori H.,
RA Hosokawa S., Masukawa M., Arikawa K., Chiden Y., Hayaishi M.,
RA Okamoto M., Ando T., Aoki H., Arita K., Hamada M., Harada C.,
RA Hijishita S., Honda M., Ichikawa Y., Idenuma A., Iijima A., Kamiya K.,
RA Ikono M., Itoh S., Itoh T., Itoh Y., Iwabuchi A., Kamei K.,
RA Karasawa W., Katagiri S., Kikuta A., Kobayashi N., Kono I.,
RA Machita K., Maehara T., Mizuno H., Mizubayashi T., Mukai Y.,
RA Nagasaki H., Nakashima M., Nakama Y., Nakamichi Y., Nakamura M.,
RA Namiki N., Negishi M., Ohta I., Ono N., Saji S., Sakai, K. Shibata M.,
RA Shinokawa T., Shomura A., Song J., Takazaki Y., Terasawa K., Tsuji K.,
RA Waki K., Yamagata H., Yamane H., Yoshiki S., Yoshihara R., Yukawa K.,
RA Zhong H., Iwama H., Endo T., Ito H., Hahn J.H., Kim H.I., Eun M.Y.,
RA Yano M., Jiang J., Gojobori T.;
RT "The genome sequence and structure of rice chromosome 1.";
RL Nature 420:312-316(2002).
DR EMBL; AP003284; BAB91742.1; -.
DR HSSP; F17679; IGFN.
DR Gramene; QBLR32; -.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003700; P:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR007087; Znf_C2H2.
DR InterPro; IPR000679; Znf_GATA.
DR Pfam; PF00320; GATA; 1.
DR SMART; SM00401; Znf_GATA; 1.
DR PROSITE; PS01114; GATA_ZNF_FINGER_2; 1.
DR PROSITE; PS00028; ZINC_FINGER_C2H2; 1; UNKNOWN 1.
SQ SEQUENCE 242 AA; 24923 MW; EPE29SD0A8F8628E CRC64;

Query Match 57.9%; Score 44; DB 2; Length 242;
Best Local Similarity 52.9%; Pred. No. 42;
Matches 9; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 KKKRKAANAALLPVLLA 17
Db 155 KEERRAANAALLPVLLA 171

```

```

RESULT 4
QBL4U4 PRELIMINARY; PRT; 802 AA.
AC QBL4U4;
DT 01-OCT-2002 (TREMELrel. 22, Created)
DT 01-OCT-2002 (TREMELrel. 22, Last sequence update)
DT 05-JUL-2004 (TREMELrel. 27, Last annotation update)
DE Putative receptor-like protein kinase (Hypothetical protein
DE OSJNB0038A07.18).
GN ORFNames=OSJNB0038A07.18, OSJNB0038A07.18;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA Buell C.R., Yuan Q., Ouyang S., Liu J., Gansberger K., Kim M.M.,
RA Overton II L.L., Bera J.J., Tsitrin T., Krol M.I., Jarrahi B.B.,
RA Jin S.S., Koo H., Ziemann V., Hsiao J., Blunt S., Vanaken S.S.,
RA Utterback T.T., Feldblyum T.V., Yang Q.Q., Haas B.J., Suh B.B.,
RA Peterson J.J., Quackenbush J., White O., Salzberg S.L., Fraser C.M.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Buell R.;
RN Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Buell C.R., Yuan Q., Ouyang S., Liu J., Gansberger K., Kim M.M.,
RA Overton II L.L., Bera J.J., Tsitrin T., Krol M.I., Jarrahi B.B.,
RA Jin S.S., Koo H., Ziemann V., Hsiao J., Blunt S., Vanaken S.S.,
RA Utterback T.T., Feldblyum T.V., Yang Q.Q., Haas B.J., Suh B.B.,
RA Peterson J.J., Quackenbush J., White O., Salzberg S.L., Fraser C.M.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Buell R.;
RN Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RA The Rice Chromosome 10 Sequencing Consortium;
RT "In-depth view of structure, activity, and evolution of rice
RT chromosome 10.";
RL Science 300:1566-1569(2003).
RN [6]
RP SEQUENCE FROM N.A.
RA Buell C.R., Wing R.A., McCombie W.R., Messing J., Yuan Q.;
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC092389; AAM88626.1; -.
DR EMBL; AC113948; AAM94518.1; -.
DR EMBL; AB017114; AAP54775.1; -.
DR Gramene; QBL4U4; -.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004672; F:protein kinase activity; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR011009; Kinase_like.
DR InterPro; IPR001611; LRR.
DR InterPro; IPR007090; LRR_plant.
DR Pfam; PF00560; LRR_1; 9.
DR PRINTS; PR00019; LEURICHRPT.
DR PRODOM; PD000001; Prot_kinase; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
KW Hypothetical protein; Kinase; Receptor.
SQ SEQUENCE 802 AA; 85525 MW; 41A650742C4C6144 CRC64;

Query Match 57.9%; Score 44; DB 2; Length 802;
Best Local Similarity 56.2%; Pred. No. 1.3e+02;
Matches 9; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

```

```
QY 1 KKKKRAAAVLLPVLL 16
   :: : ||||| : |||||
Db 2 RRSFRAAAAVLLPVLL 17

RESULT 5
PPBT_CHICK STANDARD; PRT; 519 AA.
AC Q92058;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Alkaline phosphatase, tissue-nonspecific isozyme precursor
DE (EC 3.1.3.1) (AP-TNAP) (liver/bone/kidney isozyme).
GN Name=ALPL;
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9610751; PubMed=8563025;
RA Crawford K., Weissig H., Binette F., Millan J.L., Goetinck P.F.;
RT "Tissue-nonspecific alkaline phosphatase participates in the
RT establishment and growth of feather germs in embryonic chick skin
RT cultures.";
RL Dev. Dyn. 204:48-56 (1995).
RC -I- CATALYTIC ACTIVITY: An orthophosphoric monoester + H(2)O = an
CC alcohol + phosphate.
CC -I- SUBUNIT: Homodimer (By similarity).
CC -I- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor (By
CC similarity).
CC -I- SIMILARITY: Belongs to the alkaline phosphatase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
-----
CC EMBL; U9108; AAA92562.1; -.
CC HSSP; P05187; 1EW2.
CC InterPro; IPR001952; Alk_phosphatse.
CC Pfam; PF00245; Alk_phosphatase; 1.
CC PRINTS; PR00113; ALKPHPTASE.
CC SMART; SM00098; alkPPC; 1.
CC PROSITE; PS00123; ALKALINE_PHOSPHATASE; 1.
KW Glycoprotein; GPI-anchor; Hydrolase; Lipoprotein; Magnesium; Membrane;
KW Multigene family; Phosphorylation; Signal; Zinc.
FT SIGNAL 1 16 Potential.
FT CHAIN 17 498 Alkaline phosphatase, tissue-nonspecific
FT isozyme.
FT PROPEP 499 519 Removed in mature form (Potential).
FT ACT_SITE 109 109 Phosphoserine intermediate (By
FT similarity).
FT FT
FT LPID 498 498 GPI-anchor amidated serine (Potential).
FT CARBOHYD 139 139 N-linked (GLCNAC...) (Potential).
FT CARBOHYD 229 229 N-linked (GLCNAC...) (Potential).
FT CARBOHYD 278 278 N-linked (GLCNAC...) (Potential).
FT CARBOHYD 302 302 N-linked (GLCNAC...) (Potential).
FT CARBOHYD 429 429 N-linked (GLCNAC...) (Potential).
FT CARBOHYD 429 429 N-linked (GLCNAC...) (Potential).
SQ SEQUENCE 519 AA; 56760 MW; 7934C0EC3B17B89 CRC64;

Query Match 56.6%; Score 43; DB 1; Length 519;
Best Local Similarity 76.9%; Pred. No. 1.2e+02;
Matches 10; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 4 RKAATAVLLPVLL 16
   | | | | |

```

```
Db 501 RPAATATLLPVLL 513

RESULT 6
Q99PG1 PRELIMINARY; PRT; 582 AA.
AC Q99PG1;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Matrix metalloprotease MT1-MMP.
GN Name=MMPL4;
OS Cricetulus griseus (Chinese hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Cricetulus.
OX NCBI_TaxID=10029;
RN [1]
RP SEQUENCE FROM N.A.
RA Borrelli-Pages M., Arribas J.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF308662; AAG34676.2; -.
DR HSSP; P50281; 1BQQ.
DR MEROPS; M10.014; -.
DR GO; GO:0005578; C:extracellular matrix (sensu Metazoa); IEA.
DR GO; GO:0004222; F:metalloendopeptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR000585; Hemopexin.
DR InterPro; IPR006026; Peptidase M.
DR InterPro; IPR001818; Pept M10A M12B.
DR InterPro; IPR006025; Pept_M_Zn_BS.
DR InterPro; IPR009070; PGSD_like.
DR Pfam; PF000445; Hemopexin; 4.
DR Pfam; PF03933; Peptidase M10_N; 1.
DR PRINTS; PR00138; MATRIXIN.
DR SMART; SM00120; HX; 4.
DR SMART; SM00235; Zmc; 1.
DR PROSITE; PS00546; CYSTEINE_SWITCH; 1.
DR PROSITE; PS00024; HEMOPEXIN; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.
KW Metalloprotease; Protease.
SQ SEQUENCE 582 AA; 66068 MW; 390EC6AA632D32DF8 CRC64;

Query Match 56.6%; Score 43; DB 2; Length 582;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 AAVLLPVLL 16
   |||||
Db 539 AAVLLPVLL 548

RESULT 7
Q6ND81 PRELIMINARY; PRT; 133 AA.
AC Q6ND81;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein precursor.
GN OrderedLocNames=RPA0228;
OS Rhodopsuomonas palustris.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiales; Rhodopsuomonas.
OX NCBI_TaxID=1076;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CGA009 / ATCC BAA-98;
RX PubMed=14704707; DOI=10.1038/nbt993;
RA Larimer F.W., Chain P., Hauser L., Lamerdin J.E., Malfatti S., Do L.,
RA Land M.L., Pelletier D.A., Beatty J.T., Lang A.S., Tabita F.R.,
RA Gibson J.L., Hanson T.E., Bobst C., Torres y Torres J.L., Peres C.,
RA Harrison F.H., Gibson J., Harwood C.S.;
```

RT "Complete genome sequence of the metabolically versatile  
 RT Photosynthetic bacterium Rhodospirillum rubrum  
 RL Nat. Biotechnol. 22:55-61(2004).  
 DR EMBL; BX572593; CAE25672.1; -  
 KW Complete proteome; Hypothetical protein; Signal.  
 FT SIGNAL 1 37  
 SQ SEQUENCE 133 AA; 14481 MW; E22F622199718B5B CRC64;

Query Match 55.3%; Score 42; DB 2; Length 133;  
 Best Local Similarity 76.9%; Pred. No. 52;  
 Matches 10; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 5 KAAAVLLPVLLA 17  
 DB 12 KSTAAALLPVLLA 24  
 | : || |||||  
 K A A V L L P V L L A K S T A A A L L P V L L A

RESULT 8  
 Q9CR24 PRELIMINARY; PRT; 210 AA.

AC Q9CR24;  
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)  
 DE Mus.musculus adult male tongue cDNA, RIKEN full-length enriched  
 DE library, clone:2310039H17 product:hypothetical NUDIX hydrolase  
 DE containing protein, full insert sequence (Mus musculus adult male  
 DE kidney cDNA, RIKEN full-length enriched library, clone:0610012H14  
 DE product:hypothetical NUDIX hydrolase containing protein, full insert  
 DE sequence)(Mus musculus 9.5 days embryo parthenogenote cDNA, RIKEN  
 DE full-length enriched library, clone:B130046A04 product:hypothetical  
 DE NUDIX hydrolase containing protein, full insert sequence)(Nudix  
 DE (Nucleoside diphosphate linked moiety X)-type motif 8).  
 GN Name=Nudt8;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10990;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Kidney, Parthenogenote, and Tongue;  
 RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;  
 RA Carninci P., Hayashizaki Y.;  
 RL Meth. Enzymol. 303:19-44(1999).  
 RN [2]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Kidney, Parthenogenote, and Tongue;  
 RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;  
 RA RIKEN FANTOM Consortium;  
 RT "Functional annotation of a full-length mouse cDNA collection."  
 RL Nature 409:685-690(2001).  
 RN [3]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Kidney, Parthenogenote, and Tongue;  
 RA The FANTOM Consortium,  
 RT "Analysis of the mouse transcriptome based on functional annotation of  
 RT 60,770 full-length cDNAs."  
 RL Nature 420:563-573(2002).  
 RN [4]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Kidney, Parthenogenote, and Tongue;  
 RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;  
 RA Carninci P., Shibata Y., Hayatsu N., Sugahara I., Shibata K., Itoh M.,  
 RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;  
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to  
 RT prepare full-length cDNA libraries for rapid discovery of new genes."  
 RL Genome Res. 10:1617-1630(2000).  
 RN [5]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Kidney, Parthenogenote, and Tongue;  
 RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;

RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,  
 RA Konno H., Akiyama J., Nishi K., Kitsuami T., Tashiro H., Itoh M.,  
 RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,  
 RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,  
 RA Fujiwara S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,  
 RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,  
 RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;  
 RT "RIKEN integrated sequence analysis (RISA) system-384-format  
 RT sequencing pipeline with 384 multiplexed capillary sequencer."  
 RL Genome Res. 10:1757-1771(2000).  
 RN [6]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Kidney, and Tongue;  
 RA Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H.,  
 RA Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,  
 RA Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,  
 RA Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,  
 RA Kawai J., Kojima Y., Konno H., Kouda M., Koya S., Kurihara C.,  
 RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,  
 RA Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,  
 RA Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,  
 RA Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,  
 RA Tejima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,  
 RA Muramatsu M., Hayashizaki Y.;  
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
 RN [7]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Parthenogenote;  
 RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,  
 RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,  
 RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,  
 RA Hori F., Imotani K., Ishii Y., Itoh M., Kasukawa T.,  
 RA Kato H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,  
 RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,  
 RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,  
 RA Saito R., Saito H., Sakai C., Sakai K., Sakazume N., Sano H.,  
 RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,  
 RA Tagawa A., Takahashi F., Takaku-Akashira S., Takeda Y., Tanaka T.,  
 RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;  
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
 RN [8]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Brain;  
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hong L.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hsieh F.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Lequellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettner M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young J.W., Green E.D., Dickson M.C.,  
 RA Blakesley R.W., Touchman J.W., Shevchenko Y., Bouffard G.G.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 RA Krzywinski M.I., Skalska U., Smallos D.E., Schnerch A., Schein J.E.,  
 RA Jones S.J., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences."  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [9]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Brain;  
 RA Strausberg R.;  
 RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AK009700; BAB26447.1; -  
 DR EMBL; AK002605; BAB2224.1; -  
 DR EMBL; AK045197; BAC32259.1; -  
 DR EMBL; BC056443; AAB56443.1; -

```

DR MGD; MGI:1913637; Nudr8.
DR CO; CO:0005739; C-mitochondrion; IDA.
DR InterPro; IPR000086; NUDIX_hydrolase.
DR Pfam; PF00293; NUDIX; 1.
DR KQ Hydrolase; Hypothetical protein.
SQ SEQUENCE 210 AA; 23253 MW; 2E153033C466A6FE CRC64;

Query Match 55.3%; Score 42; DB 2; Length 210;
Best Local Similarity 60.0%; Pred. No. 80;
Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 2 KKRKAAAVALLPVLL 16
| | | | | | | | | | | | | | | |
Db 26 RSRPAAAAVILPVLCL 40

RESULT 9
Q72PH8 PRELIMINARY; PRT; 221 AA.
AC Q72PH8;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Putative lipoprotein.
GN OrderedLocusNames=LI1C12493;
OS Leptospira interrogans (serogroup Icterohaemorrhagiae / serovar Copenhageni).
OC Bacteria; Spirochaetes; Spirochaetales; Leptospiraceae; Leptospira.
OX NCBI_TaxID=44275;
RN [1]_
RP SEQUENCE FROM N.A.
RC STRAIN=Fiocruz LI-130;
RX PubMed=15028702; DOI=10.1128/JB.186.7.2164-2172.2004;
RA Nascimento A.L.T.O., Ko A.I., Martins E.A.L., Monteiro-Vitorello C.B., Ho P.L., Haake D.A., Vertovski-Almeida S., Hartskeerl R.A., Marques M.V., Oliveira M.C., Mencia C.F.M., Leite L.C.C., Carver H., Coutinho L.L., Degraeve W.M., Dellagostin O.A., El-Dorry H., Ferro E.S., Ferro M.I.T., Furlan L.R., Gamberini M., Gigliotti E.A., Goes-Neto A., Goldman G.H., Goldman M.H.S., Harakava R., Jeronimo S.M.B., Junqueira-de-Azevedo I.L.M., Kimura E.T., Kuramae E.E., Lemos E.G.M., Lemos M.V.F., Marino C.L., Nunes L.R., de Oliveira R.C., Pereira G.G., Reis M.S., Schriefer A., Siqueira W.J., Sommer P., Tsai S.M., Simpson A.J.G., Ferro J.A., Camargo L.E.A., Katajima J.P., Setubal J.C., Van Sluys M.A.;
RT "Comparative genomics of two Leptospira interrogans serovars reveals novel insights into physiology and pathogenesis.";
RL J. Bacteriol. 186:2164-2172(2004).
DR EMBL; AB017297; AAS71058.1; -.
KW Complete proteome.
SQ SEQUENCE 221 AA; 25241 MW; 9119A0A728BCE258 CRC64;

Query Match 55.3%; Score 42; DB 2; Length 221;
Best Local Similarity 53.3%; Pred. No. 83;
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 2 KKRKAAAVALLPVLL 16
| | | | | | | | | | | | | | | |
Db 25 KFRKAIVAILMPMIL 39

RESULT 10
Q8F6U9 PRELIMINARY; PRT; 221 AA.
AC Q8F6U9;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=LAI202;
OS Leptospira interrogans.
OC Bacteria; Spirochaetes; Spirochaetales; Leptospiraceae; Leptospira.
OX NCBI_TaxID=173;
RN [1]_
RP SEQUENCE FROM N.A.
RC STRAIN=56601 / Serogroup Icterohaemorrhagiae / Serovar lai;
RX MEDLINE=22598143; PubMed=12712204; DOI=10.1038/nature01597;
RA Ren S.-X., Fu G., Jiang X.-G., Zeng R., Miao Y.-G., Xu H., Zhang Y.-X., Xiong H., Lu G., Lu L.-F., Jiang H.-Q., Jia J., Tu Y.-F., Jia J.-X., Gu W.-Y., Zhang Y.-Q., Cai Z., Sheng H.-H., Yin H.-F., Zhang Y., Zhu G.-P., Wan M., Huang H.-L., Qian Z., Wang S.-Y., Ma W., Yao Z.-J., Shen Y., Qiang B.-Q., Xia Q.-C., Guo X.-K., Danchin A., Saint Glrons I., Somerville R.L., Wen Y.-M., Shi M.-H., Chen Z., Xu J.-G., Zhao G.-P.;
RT "Unique physiological and pathogenic features of Leptospira RT interrogans revealed by whole-genome sequencing.";
RL Nature 422:888-893(2003).
DR EMBL; AB011302; AAN48400.1; -.
KW Complete proteome.
SQ SEQUENCE 221 AA; 25241 MW; 9119A0A728BCE258 CRC64;

Query Match 55.3%; Score 42; DB 2; Length 221;
Best Local Similarity 53.3%; Pred. No. 83;
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 2 KKRKAAAVALLPVLL 16
| | | | | | | | | | | | | | | |
Db 25 KFRKAIVAILMPMIL 39

RESULT 11
Q8G6L8 PRELIMINARY; PRT; 268 AA.
AC Q8G6L8;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=BL0621;
OS Bifidobacterium longum.
OC Bacteria; Actinobacteria; Actinobacteridae; Bifidobacteriales; Bifidobacteriaceae; Bifidobacterium.
OX NCBI_TaxID=216816;
RN [1]_
RP SEQUENCE FROM N.A.
RC STRAIN=NCC 2705;
RX MEDLINE=22294977; PubMed=12381787; DOI=10.1073/pnas.212527599;
RA Schell M.A., Karmirantzou M., Snel B., Vilanova D., Berger B., Pessi G., Zwaalen M.-C., Desiere F., Bork P., Delley M., Fridmore R.D., Arigoni F.;
RT "The genome sequence of Bifidobacterium longum reflects its adaptation to the human gastrointestinal tract.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:14422-14427(2002).
DR EMBL; AB014684; AAN24444.1; -.
DR InterPro; IPR002725; DUF45.
DR InterPro; IPR006025; Pept_M_Zn_BS.
DR Pfam; PF01863; DUF45; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 268 AA; 29679 MW; 04F91A150C29085E CRC64;

Query Match 55.3%; Score 42; DB 2; Length 268;
Best Local Similarity 57.9%; Pred. No. 99;
Matches 11; Conservative 4; Mismatches 2; Indels 2; Gaps 1;

QY 1 KKRKAAAVALLPVLL 17
| | | | | | | | | | | | | | | |
Db 160 ERKQAAAAINAALPOLLA 178

RESULT 12
Q8FYJ9 PRELIMINARY; PRT; 293 AA.
AC Q8FYJ9;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)

```

DE Membrane protein, putative.  
 GN OrderedLocNames=BR1873;  
 OS Brucella suis.  
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
 OC Brucellaceae; Brucella.  
 OX NCBI\_TaxID=29451;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RC STRAIN=1330 / Biovar 1;  
 RX MEDLINE=22247741; PubMed=12271122; DOI=10.1073/pnas.192319099;  
 RA Paulsen I.T., Seshadri R., Nelson K.E., Eisen J.A., Heidelberg J.F.,  
 RA Read I.D., Dodson R.T., Umayam L.A., Brinkac L.M., Beanan M.J.,  
 RA Daugherty S.C., DeBoy R.T., Durkin A.S., Kolonay J.F., Madupu R.,  
 RA Nelson W.C., Ayodeji B., Kraul M., Shetty J., Malek J.A.,  
 RA Van Aken S.E., Riedmuller S., Tetelini H., Gill S.R., White O.,  
 RA Salzberg S.L., Hoover D.L., Lindler L.E., Halling S.M., Boyle S.M.,  
 RA Fraser C.M.;  
 RT "The Brucella suis genome reveals fundamental similarities between  
 animal and plant pathogens and symbionts.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:13148-13153 (2002).  
 DR EMBL; AE014478; AAN30768.1; --  
 DR TIGR; BR1873; --  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR InterPro; IPR000620; DUF6.  
 DR Pfam; PF00892; DUF6; 1.  
 KW Complete proteome.  
 SQ SEQUENCE 293 AA; 30733 MW; 4DFD4347112050AE CRC64;

Query Match 55.3%; Score 42; DB 2; Length 293;  
 Best Local Similarity 81.8%; Pred. No. 1.1e+02;  
 Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 7 AAALLPVL 17  
 DB 156 AAALLPVL 166  
 |||:|||||

## RESULT 13

QY9J99  
 ID Q8VJ99 PRELIMINARY; PRT; 297 AA.  
 AC Q8VJ99;  
 DT 01-WAR-2002 (TRENBLrel. 20, Created)  
 DT 01-WAR-2002 (TRENBLrel. 20, Last sequence update)  
 DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)  
 DE TRANSPORTER, DME FAMILY.  
 GN OrderedLocNames=BMEI0187;  
 OS Brucella melitensis.  
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;  
 OC Brucellaceae; Brucella.  
 OX NCBI\_TaxID=29459;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=16M / ATCC 23456 / Biotype 1;  
 RX MEDLINE=20020109; PubMed=11756688; DOI=10.1073/pnas.221575398;  
 RA DelVecchio V.G., Kapur R.J., Redkar R.J., Patra G., Mujer C., Los T.,  
 RA Ivanova N., Anderson I., Bhattacharya A., Lykdis A., Reznik G.,  
 RA Jablonks L., Larsen N., D'Souza M., Bernal A., Mazur M., Goldsman E.,  
 RA Selkov E., Elzer P.H., Hagius S., O'Callaghan D., Letesson J.-J.,  
 RA Haskorn R., Kyrides N.C., Overbeek R.;  
 RT "The genome sequence of the facultative intracellular pathogen  
 Brucella melitensis.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:443-448 (2002).  
 DR EMBL; AE009460; AAL51369.1; --  
 DR PIR; AF3275; AF3275.  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR Pfam; PF00892; DUF6; 1.  
 KW Complete proteome.  
 SQ SEQUENCE 297 AA; 31306 MW; E8CA2D3CEAB94B85 CRC64;

Query Match 55.3%; Score 42; DB 2; Length 297;  
 Best Local Similarity 81.8%; Pred. No. 1.1e+02;  
 Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 7 AAALLPVL 17  
 DB 160 AAALLPVL 170  
 |||:|||||

## RESULT 14

Q8KO12  
 ID Q8KO12 PRELIMINARY; PRT; 524 AA.  
 AC Q8KO12;  
 DT 01-OCT-2002 (TRENBLrel. 22, Created)  
 DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)  
 DT 01-WAR-2004 (TRENBLrel. 26, Last annotation update)  
 DE 4930403J22Rik protein.  
 GN Name=4930403J22Rik;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=mix FVB/N; TISSUE=Mammary tumor;  
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Schuler G.D.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.P., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Frange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettner M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 RA Krzywinski M.I., Skalek U., Smalls D.E., Schnerch A., Schein J.E.,  
 RA Jones S.J., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16999-16903 (2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=mix FVB/N; TISSUE=Mammary tumor;  
 RA Strausberg R.;  
 RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC031368; AAH31368.1; --  
 DR MGD; MGI:1921050; 4930403J22Rik.  
 DR InterPro; IPR001395; Aldo/ket red.  
 DR InterPro; IPR008940; Prenyl\_trans.  
 DR InterPro; IPR001440; TPR.  
 DR Pfam; PF00515; TPR\_1; 4.  
 DR SMART; SM00028; TPR; 7.  
 DR PROSITE; PS00063; ALDOXETO\_REDUCTASE\_3; UNKNOWN 1.  
 DR PROSITE; PS00005; TPR; 5.  
 DR PROSITE; PS02093; TPR\_REGION; 1.  
 KW Repeat; TPR repeat.  
 SQ SEQUENCE 524 AA; 59060 MW; 27A07F39195F2F0D CRC64;

Query Match 55.3%; Score 42; DB 2; Length 524;  
 Best Local Similarity 53.3%; Pred. No. 1.8e+02;  
 Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 2 KKKKAAAVLLPVL 16  
 DB 220 KKKKPAAILGILL 234  
 |||:|||||

## RESULT 15

PTIB\_BACSU  
 ID PTIB\_BACSU STANDARD; PRT; 527 AA.  
 AC P54715;

DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 25-OCT-2004 (Rel. 45, Last annotation update)  
 DE PTS system, arbutin-like IIBC component (Phosphotransferase enzyme II,  
 DE BC component) (SC 2.7.1.69).  
 GN Name=glvC; Synonyms=glv-2, glvCB; OrderedLocusNames=BSU08200;  
 OS *Bacillus subtilis*.  
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
 OX NCBI\_TaxID=1423;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=168;  
 RX MEDLINE=96262713; PubMed=8704981;  
 RA Yamamoto H., Uchiyama S., Fajar A.N., Ogasawara N., Sekiguchi J.:  
 RT "Determination of a 12 kb nucleotide sequence around the 76 degrees  
 RT region of the *Bacillus subtilis* chromosome.";  
 RL Microbiology 142:1417-1421(1996).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=168;  
 RX MEDLINE=98044033; PubMed=9384377; DOI=10.1038/36786;  
 RA Kunst F., Ogasawara N., Mosser I., Albertini A.M., Alloni G.,  
 RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,  
 RA Brouillet S., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,  
 RA Choi S.K., Codani J.J., Conerton I.F., Cummings N.J., Daniel R.A.,  
 RA Denicot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,  
 RA Entian K.-D., Errington J., Fabret C., Ferrari E., Foulger D.,  
 RA Fritz C., Fujita M., Fujita Y., Funa S., Galizzi A., Galleron N.,  
 RA Gim S.Y., Glaeser P., Goffeau A., Golightly E.J., Grandi G.,  
 RA Guisepi G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,  
 RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M.,  
 RA Jones L.-M., Joris B., Karamata D., Kasahara Y., Klaert-Blanchard M.,  
 RA Klein C., Kobayashi Y., Koetter P., Koningsstein G., Krogh J.,  
 RA Kumano M., Kurita K., Lapidus A., Lardinois S., Lauber J.,  
 RA Lazarevic V., Lee S.M., Levine A., Liu H., Masuda S., Mauel C.,  
 RA Medigue C., Medina M., Mellado R.P., Mizuno M., Moestl D., Nakai S.,  
 RA Noback M., Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B.,  
 RA Park S.H., Parro V., Pohl T.M., Portetelle D., Porwollik S.,  
 RA Prescott A.M., Presecan E., Pujic P., Purnelle B., Rapoport G.,  
 RA Rey M., Reynolds S., Rieger M., Rivolta C., Rocha E., Roche B.,  
 RA Rose M., Sadaie Y., Sato T., Scanlan E., Schleich S., Schroeter R.,  
 RA Scoffone F., Sekiguchi J., Sekowska A., Seror S.J., Seror P.,  
 RA Shin B.S., Soldo B., Sorokin A., Tacconi E., Takagi T., Takahashi H.,  
 RA Takenaru K., Takeuchi M., Tamakoshi A., Tanaka T., Terpetra P.,  
 RA Tognoni A., Tosato V., Uchiyama S., Vandenbol M., Vannier P.,  
 RA Vassarotti A., Viari A., Wambutt R., Wedler E., Wedler H.,  
 RA Weitzenecker T., Winters P., Wipat A., Yamamoto H., Yamane K.,  
 RA Yasumoto K., Yata K., Yoshida K., Yoshikawa H.F., Zumstein E.,  
 RA Yoshikawa H., Danchin A.;  
 RT "The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.";  
 RL Nature 390:249-256(1997).  
 RN [3]  
 RP INDUCTION.  
 RX MEDLINE=21392183; PubMed=1149864;  
 RX DOI=10.1128/JB.183.17.5110-5121.2001;  
 RA Yamamoto H., Serizawa M., Thompson J., Sekiguchi J.;  
 RT "Regulation of the glv operon in *Bacillus subtilis*: YfiA (GlvR) is a  
 RT positive regulator of the operon that is repressed through CcpA and  
 RT cre.";  
 RL J. Bacteriol. 183:5110-5121(2001).  
 CC -1- FUNCTION: This is a component of the phosphoenolpyruvate-dependent  
 CC sugar phosphotransferase system (PTS), a major carbohydrate active  
 CC -transport system. The IICD domains contain the sugar binding site  
 CC and the transmembrane channel; the IIA domain contains the primary  
 CC phosphorylation site (the donor is phospho-HPr); IIA transfers its  
 CC phosphoryl group to the IIB domain which finally transfers it to  
 CC the sugar.  
 CC -1- CATALYTIC ACTIVITY: Protein N-phosphohistidine + sugar = protein  
 CC histidine + sugar phosphate.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Probable).  
 CC -1- INDUCTION: By maltose; repressed by glucose.

CC -1- SIMILARITY: Contains 1 PTS EIIB domain.  
 CC -1- SIMILARITY: Contains 1 PTS EIIC domain.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; D50543; BAA09105.1; -;  
 CC EMBL; Z99108; CAB12649.1; -;  
 CC PIR; G69635; G69635.  
 CC HSP; P05053; IIBA.  
 CC Subtilist; BG11848; glvC.  
 CC InterPro; IPR001996; Ptrans EIIB.  
 CC InterPro; IPR003352; Ptrans EIIC.  
 CC InterPro; IPR010975; PTS IIBC alpha.  
 CC InterPro; IPR004719; PTS IIC glc.  
 CC Pfam; PF00367; PTS EIIB; 1.  
 CC Pfam; PF02378; PTS EIIC; 1.  
 CC ProDom; PD001476; Ptrans EIIB; 1.  
 CC TIGRPFAMs; TIGR00826; EIIB glc; 1.  
 CC TIGRPFAMs; TIGR00852; pts-glc; 1.  
 CC TIGRPFAMs; TIGR02002; PTS-II-BC-glcB; 1.  
 CC TIGRPFAMs; TIGR01998; PTS-II-BC-nag; 1.  
 CC TIGRPFAMs; TIGR02005; PTS-IIBC-alpha; 1.  
 CC PROSITE; PS01035; PTS EIIB CYS; 1.  
 CC Complete proteome; Phosphorylation; Phosphotransferase system;  
 KW Sugar transport; Transferase; Transmembrane; Transport.  
 FT DOMAIN 1 367  
 FT TRANSMEM 376 527  
 FT TRANSMEM 8 28  
 FT TRANSMEM 59 79  
 FT TRANSMEM 93 113  
 FT TRANSMEM 132 152  
 FT TRANSMEM 173 193  
 FT TRANSMEM 200 220  
 FT TRANSMEM 224 244  
 FT TRANSMEM 276 296  
 FT TRANSMEM 305 325  
 FT TRANSMEM 326 346  
 FT TRANSMEM 357 377  
 FT TRANSMEM 382 402  
 FT MOD\_RES 471 471  
 SQ SEQUENCE 527 AA; 58055 MW; D708C38048D4D63C CRC64;  
 Query Match 55.3%; Score 42; DB 1; Length 527;  
 Best Local Similarity 50.0%; Pred. No. 1.8e+02;  
 Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;  
 OY 1 KKKRKAATAAVLLPVLL 16  
 DB 300 KEKKLVAGLLIPVTL 315  
 RESULT 16  
 Q8BT03  
 ID Q8BT03 PRELIMINARY; PRT; 607 AA.  
 AC Q8BT03;  
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Mus musculus 12 days embryo embryonic body below diaphragm region  
 DE cDNA, RIKEN full-length enriched library, clone:373240C04  
 DE product:hypothetical Aldo/keto reductase family containing protein,  
 DE full insert sequence.  
 DE Mus musculus (Mouse)  
 OS Mus musculus (Mouse)  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]





```

RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA Kato H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.,
RL Submitted (Apr-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK082526; BAC38520.1; -.
DR MGD; MGI:1921050; 4930403J2R1K.
DR InterPro; IPR001395; Aldo/ket red.
DR InterPro; IPR008940; Prenyl_trans.
DR InterPro; IPR001440; TPR.
DR Pfam; PF00515; TPR 1; 4.
DR SMART; SM00028; TPR; 7.
DR PROSITE; PS00063; ALDO/KETO_REDUCTASE_3; UNKNOWN_1.
DR PROSITE; PS50005; TPR; 5.
DR PROSITE; PS50293; TPR_REGION; 1.
KW Hypothetical protein; Repeat; TPR repeat.
SQ SEQUENCE 621 AA; 69489 MW; 6BLBAD7319D2B1A6 CRC64;

Query Match 55.3%; Score 42; DB 2; Length 621;
Best Local Similarity 53.3%; Pred. No. 2.1e+02;
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 2 KKKKAAVLLPVLL 16
||:| ||::|:|
Db 317 KKKKPVAAIILGILL 331

RESULT 18
Q891J7 PRELIMINARY; PRT; 628 AA.
AC Q891J7;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Blr5637 protein.
DE OrderedLocusNames=blr5637;
OS Bradyrhizobium japonicum.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Bradyrhizobium.
OX NCBI_TaxID=375;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=USDA110;
RX MEDLINE=22484998; PubMed=12597275;
RA Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiumi T.,
RA Sasamoto S., Watanabe A., Idesawa K., Iriguchi M., Kawashima K.,
RA Kohara M., Matsumoto M., Shimo S., Teurouka H., Wada T., Yamada M.,
RA Tabata S.;
RA "Complete genomic sequence of nitrogen-fixing symbiotic bacterium
RT Bradyrhizobium japonicum USDA110.";
RL DNA Res. 9:189-197(2002).
DR EMBL; AR005955; BAC50902.1; -.
DR GO; GO:0004383; F:guanylate cyclase activity; IEA.
DR GO; GO:0016829; P:lyase activity; IEA.
DR GO; GO:0007242; P:intracellular signaling cascade; IEA.
DR InterPro; IPR007890; CHASE2.
DR InterPro; IPR001054; G_cyclase.
DR Pfam; PF05226; CHASE2_1.
DR SMART; SM00211; Guanylate_cyc; 1.
DR PROSITE; PS00044; CYCC; 1.
KW Complete proteome; Lyase.
SQ SEQUENCE 628 AA; 66669 MW; 162B19FB9DAA9342 CRC64;

Query Match 55.3%; Score 42; DB 2; Length 628;
Best Local Similarity 90.9%; Pred. No. 2.2e+02;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 6 AAAAVLLPVLL 16

```

```

Db 317 ALAAVLLPVLL 327
| |||||
Q8BG19 PRELIMINARY; PRT; 741 AA.
ID Q8BG19;
AC Q8BG19;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Mus musculus adult male diencephalon cDNA, RIKEN full-length enriched
DE library, clone:9330126C21 product:hypothetical Aldo/keto reductase
DE family containing protein, full insert sequence (Mus musculus 7 days
DE neonate cerebellum cDNA, RIKEN full-length enriched library,
DE clone:A730081108 product:hypothetical Aldo/keto reductase family
DE containing protein, full insert sequence) (Mus musculus adult retina
DE cDNA, RIKEN full-length enriched library, clone:A930010K05
DE product:hypothetical Aldo/keto reductase family containing protein,
DE full insert sequence) (Mus musculus 9.5 days embryo parthenogenote,
DE cDNA, RIKEN full-length enriched library, clone:B130054J16
DE product:hypothetical Aldo/keto reductase family containing protein,
DE full insert sequence) (Mus musculus 12 days embryo spinal cord cDNA,
DE RIKEN full-length enriched library, clone:C530046F07
DE product:hypothetical Aldo/keto reductase family containing protein,
DE full insert sequence).
DE Name=4930403J2R1K;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RC TISSUE=Cerebellum, Diencephalon, Parthenogenote, Retina, and
RC Spinal cord;
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RA "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RC TISSUE=Cerebellum, Diencephalon, Parthenogenote, Retina, and
RC Spinal cord;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/350555500;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RC TISSUE=Cerebellum, Diencephalon, Parthenogenote, Retina, and
RC Spinal cord;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RC TISSUE=Cerebellum, Diencephalon, Parthenogenote, Retina, and
RC Spinal cord;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;

```

```

RC TISSUE=Cerebellum, Diencephalon, Parthenogenote, Retina, and
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Onoda Y., Ishikawa T., Ozawa K., Tanaka T., Matsushima S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.,
RA "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771 (2000).
RN [16]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J;
RC TISSUE=Cerebellum, Diencephalon, Parthenogenote, Retina, and
RC Spinal cord;
RA Fukuchi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA Adachi S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
RA Hori F., Inotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK033947; BAC28522.1; -.
DR EMBL; AK043293; BAC31518.1; -.
DR EMBL; AK044391; BAC31900.1; -.
DR EMBL; AK049748; BAC33902.1; -.
DR EMBL; AK045271; BAC32288.1; -.
DR MGD; MGI:1921050; 4930403J22Rik.
DR InterPro; IPR001395; Aldo/ket red.
DR InterPro; IPR008940; Prenyl_trans.
DR Pfam; PF00515; TPR 1; 4.
DR SMART; SM00028; TPR; 7.
DR PROSITE; PS00063; ALDOKETO_REDUCTASE_3; UNKNOWN 1.
DR PROSITE; PS50005; TPR; 5.
DR PROSITE; PS50293; TPR_REGION; 1.
KW Hypothetical protein, Repeat; TPR repeat.
SQ SEQUENCE 741 AA; 82962 MW; D1D3A380DFA04A7F CRC64;

Query Match 55.3%; Score 42; DB 2; Length 741;
Best Local Similarity 53.3%; Pred. No. 2.5e+02;
Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 2 KKKKAAAVLLPVLL 16
DB 437 KKKKPVAAIILGILL 451

RESULT 20
Q7S899 PRELIMINARY; PRT; 1049 AA.
AC Q7S899;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein.
GN Name=NCU06528.1;
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OR74A;
RA Galagan J.E., Calvo S.E., Borkovich K.A., Selker E.U., Read N.D.,

```

```

RA Jaffe D., FitzHugh W., Ma L.-J., Smirnov S., Purcell S., Rehman B.,
RA Elkins T., Engels R., Wang S., Nielsen C.B., Butler J., Endrizzi M.,
RA Qui D., Ianakiev P., Pedersen D., Nelson M., Washburne M.,
RA Selitrennikoff C.P., Kinsey J.A., Braun E.L., Zelter A., Schulte U.,
RA Kothe G.O., Jedd G., Mewes W., Staben C., Marcotte E., Greenberg D.,
RA Roy A., Foley K., Naylor J., Thmann N., Barrett R., Gnerre S.,
RA Kamal M., Kamyselis M., Mauceli E., Bielke C., Rudd S., Frishman D.,
RA Kryatova S., Rasmussen C., Metzberg R.L., Perkins D.D., Kroken S.,
RA Cogoni C., Macino G., Catchside D., Li W., Pratt R.J., Osmari S.A.,
RA DeSouza C.C., Glass L., Orbach M.J., Berglund J., Voelker R.,
RA Yarden O., Plamann M., Seiler S., Dunlap J., Radford A., Aramayo R.,
RA Natvig D.O., Alex L.A., Mannhaupt G., Ebbole D.J., Freitag M.,
RA Paulsen I., Sachs M.S., Lander E.S., Nussbaum C., Birren B.;
RT "The Genome Sequence of the Filamentous Fungus Neurospora crassa.";
RL Nature 0:0-0(2003).
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AABX01000239; EAA32564.1; -.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0005840; C:ribosome; IEA.
DR GO; GO:0003735; F:structural constituent of ribosome; IEA.
DR GO; GO:0006412; P:protein biosynthesis; IEA.
DR InterPro; IPR000253; FHA.
DR InterPro; IPR002143; Ribosomal L1.
DR PROSITE; PS50006; FHA_DOMAIN; 1.
KW Hypothetical protein.
SQ SEQUENCE 1049 AA; 115605 MW; 1447505CE6A3D828 CRC64;

Query Match 55.3%; Score 42; DB 2; Length 1049;
Best Local Similarity 69.2%; Pred. No. 3.5e+02;
Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 KKKKAAAVLLP 13
DB 1006 KKKKAAAVLLP 1018

RESULT 21
Q8S4W4 PRELIMINARY; PRT; 346 AA.
ID Q8S4W4;
AC Q8S4W4;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Syntaxin PM.
OS Porphyra yezoensis.
OC Eukaryota; Rhodophyta; Bangiophyceae; Bangiales; Bangiaceae; Porphyra.
OX NCBI_TaxID=2788;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21947832; PubMed=11950882;
RA Dacks J.B., Doolittle W.F.;
RT "Novel syntaxin gene sequences from Giardia, Trypanosoma and algae:
RT implications for the ancient evolution of the eukaryotic endomembrane
RT system.";
RL J. Cell Sci. 115:1635-1642(2002).
DR EMBL; AF04747; AAM12663.1; -.
DR HSSP; P32851; 1JTH.
DR InterPro; IPR010989; t-snare.
DR InterPro; IPR000727; t-snare.
DR Pfam; PF05739; SNARE; 1.
DR SMART; SM00397; t-SNARE; 1.
DR PROSITE; PS50192; t-SNARE; 1.
SQ SEQUENCE 346 AA; 36283 MW; F729BF6E8F9F65B CRC64;

Query Match 54.6%; Score 41.5; DB 2; Length 346;
Best Local Similarity 40.7%; Pred. No. 1.5e+02;
Matches 11; Conservative 3; Mismatches 2; Indels 1; Gaps 1;

QY 1 KKKRK-----AAAVALLPVLL 16
DB 261 KKKRKICCVILVIAIILIPVLI 287

```

## RESULT 22

Q6GX97 Q6GX97 PRELIMINARY; PRT; 67 AA.  
 ID Q6GX97  
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
 DE Matrix metalloproteinase 14 (fragment).  
 GN Name=Mmpl14;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=BALB/c; TISSUE=Uterus;  
 RA Ulrich R., Gerhauer I., Seeliger F., Baumgaertner W., Alldinger S.;  
 RT "Expression profile of matrix-metalloproteinases and their inhibitors  
 RT in acute and chronic demyelinating Theiler's murine  
 RT encephalomyelitis.";  
 RL Acta Neuropathol. 108:366-366(2004).  
 RN [2]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=BALB/c; TISSUE=Uterus;  
 RA Ulrich R.G., Seeliger F., Alldinger S., Baumgaertner W.;  
 RL Submitted (MAY-2004) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AY622974; AAT46405.1;  
 FT NON\_TER 1 1  
 FT TER 67 67  
 SQ SEQUENCE 67 AA; 7104 MW; 52CDDDF3957E3FBA CRC64;

Query Match 53.9%; Score 41; DB 2; Length 67;  
 Best Local Similarity 81.8%; Pred. No. 41;  
 Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 6 ARAAVLLPVLL 16  
 :|||||:|||||  
 Db 52 SAAAVVLPVLL 62

## RESULT 23

EVGL\_MOUSE STANDARD; PRT; 216 AA.  
 ID EVGL\_MOUSE  
 AC Q9D9S1;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE Hypothetical UPF0193 protein EVG1 homolog.  
 GN Names=C22orf23;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Testis;  
 RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;  
 RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,  
 RA Nikaido I., Osato N., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,  
 RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,  
 RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,  
 RA Schriml L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,  
 RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,  
 RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,  
 RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,  
 RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,  
 RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.B.,  
 RA Konoaga A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,  
 RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,  
 RA Nagashima T., Numata K., Okido T., Pavan W.J., Pertea G., Pesole G.,  
 RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,

RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,  
 RA Sandelin A., Schneider C., Semple C.A., Setou M., Shimada K.,  
 RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,  
 RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,  
 RA Wilming L.G., Wynshaw-Boris A., Yanagisawa M., Yang I., Yang L.,  
 RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,  
 RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,  
 RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,  
 RA Hara A., Hashizume W., Inotani K., Ishii Y., Itoh M., Kagawa I.,  
 RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,  
 RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,  
 RA Birney E., Hayashizaki Y.;  
 RT "Analysis of the mouse transcriptome based on functional annotation of  
 RT 60,770 full-length cDNAs.";  
 RL Nature 420:563-573(2002).  
 CC -|- SIMILARITY: Belongs to the UPF0193 (EVG1) family.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AK006539; BAB24642.1;  
 DR MGD; MGI:1920774; 1700088E04Rik.  
 DR InterPro; IPR007914; UPF0193.  
 DR Pfam; PF05250; UPF0193; 1.  
 KW Hypothetical protein.  
 SQ SEQUENCE 216 AA; 24737 MW; 5A36D524DC7A8F6B CRC64;

Query Match 53.9%; Score 41; DB 1; Length 216;  
 Best Local Similarity 57.1%; Pred. No. 1.2e+02;  
 Matches 8; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 4 RKAAAALLPVLLA 17  
 :|||:|:|:|  
 Db 76 KKAASAYLPPIA 89

## RESULT 24

Q6AYH8 PRELIMINARY; PRT; 216 AA.  
 ID Q6AYH8  
 AC Q6AYH8;  
 DT 25-OCT-2004 (TrEMBLrel. 28, Created)  
 DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)  
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)  
 DE Hypothetical protein.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RC TISSUE=Testis;  
 RX PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Krausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny K.D., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,

GN	Name=cocx3;
OS	Thraustochytrium aureum.
OG	Mitochondrion.
OC	Eukaryota; stramenopiles; Labyrinthulida; Thraustochytriidae;
OT	Thraustochytrium.
OX	NCBI_TaxID=42467;
RN	[1]
RP	SEQUENCE FROM N.A.
RA	Burger G., Lang B.F., Gray M.W.;
RL	Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
CC	-I- FUNCTION: Subunits I, II and III form the functional core of the enzyme complex (By similarity).
CC	-I- CATALYTIC ACTIVITY: 4 ferrocyclochrome c + O(2) = 4 ferricytochrome c + 2 H(2)O.
CC	-I- SIMILARITY: Belongs to the cytochrome c oxidase subunit 3 family.
DR	EMBL; AF288091; AAG23662.1; --
DR	GO; GO:0016021; C:integral to membrane; IEA.
DR	GO; GO:0005739; C:mitochondrion; IEA.
DR	GO; GO:0004129; F:cytochrome-c oxidase activity; IEA.
DR	GO; GO:0016491; F:oxidoreductase activity; IEA.
DR	GO; GO:0006118; P:electron transport; IEA.
DR	InterPro; IPRO00298; Cytc_oxdse_III.
DR	Pfam; PF00510; COX3; 1..
DR	Prodom; PD000382; Cytc_oxdse_III; 1.
DR	PROSITE; PS0253; COX3; 1.
KW	Mitochondrion; Oxidoreductase; Transmembrane.
SQ	SEQUENCE 262 AA; 29789 MW; C15163CF5C7EBCE8 CRC64;
Query Match	53.9%; Score 41; DB 2; Length 262;
Best Local Similarity	53.3%; Pred. No. 1.4e+02;
Matches	8; Conservative 5; Mismatches 2; Indels 0; Gaps 0;
QY	3 KRKAAAVLLPVLIA 17       ::::  :
Db	157 KRKASTALIITVILA 171
RESULT 27	
Q7QB91	PRELIMINARY; PRT; 297 AA.
ID	Q7QB91
AC	Q7QB91;
DT	01-MAR-2004 (TrEMBLrel. 26, Created)
DT	01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT	01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE	AgCP2447.
OS	Name=agCG45893; ORFNames=ENSANGG00000017897;
OS	Anopheles gambiae str. FES1.
OC	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC	Noptera; Endopterygota; Diptera; Nematocera; Culicoidea; Anophelinae.
OX	NCBI_TaxID=180454;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	STRAIN=PEST;
RA	Anopheles Genome Sequencing Consortium;
RL	Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
CC	-I- CAUTION: The sequence shown here is derived from an preliminary data.
CC	EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is preliminary data.
DR	EMBL; AAB0100880; EAA08668.1; --
SQ	SEQUENCE 297 AA; 30551 MW; 7D60098FFE259BB5 CRC64;
Query Match	53.9%; Score 41; DB 2; Length 297;
Best Local Similarity	60.0%; Pred. No. 1.6e+02;
Matches	9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
QY	3 KRKAAAVLLPVLIA 17         :    :
Db	172 KKKKAAILGPLLIA 186
RESULT 28	
Q56264	PRELIMINARY; PRT; 342 AA.
ID	Q56264

```

AC Q56264;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Phosphatase (Fragment).
OS Treponema denticola.
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Treponema.
OX NCBI_TaxID=158;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=35405;
RX MEDLINE=95197228; PubMed=7534273;
RA Ishihara K., Kuramitsu H.K.;
RT "Cloning and expression of a neutral phosphatase gene from Treponema
RT denticola.";
RL Infect. Immun. 63:1147-1152(1995).
DR EMBL; L25421; AAA73940.1; -.
DR InterPro; IPR008934; AcPase.VanPerase.
DR InterPro; IPR000326; Pesterase_PA_PTP.
DR Pfam; PF01569; PAP2; 1.
DR SMART; SM00014; acidPPC; 1.
FT NON_TER 342 342
SQ SEQUENCE 342 AA; 37951 MW; DAD677BE63073934 CRC64;

Query Match 53.9%; Score 41; DB 2; Length 342;
Best Local Similarity 57.1%; Pred. No. 1.8e+02;
Matches 8; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 3 KRKAANAALLPVLL 16
Db 151 KLIKIAAAILPLLV 164

RESULT 29
Q73KN6 PRELIMINARY; PRT; 346 AA.
AC Q73KN6;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE PAP2 family protein.
GN OrderedLocusNames=TDE2181;
OS Treponema denticola.
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Treponema.
OX NCBI_TaxID=158;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 35405 / DSM 14222;
RX PubMed=15064399; DOI=10.1073/pnas.0307639101;
RA Seshadri R., Myers G.S.A., Tettelin H., Eisen J.A., Heidelberg J.F.,
RA Dodson R.J., Davidson T.M., DeBoy R.T., Fouts D.E., Hatt D.H.,
RA Selengut J., Ren Q., Brinkac L.M., Madupu R., Kolonay J.F.,
RA Durkin S.A., Daugherty S.C., Shetty J., Shvartsbeyn A.,
RA Ghebreorgis E., Geer K., Teegaye G., Malek J.A., Ayodeji B.,
RA Shatman S., McLeod M.P., Snajls D., Howell J.K., Pal S., Amin A.,
RA Vashisth P., McNeill T.Z., Xiang Q., Sodergren E., Baca E.,
RA Weinstock G.M., Norris S.J., Fraser C.M., Paulsen I.T.;
RT "Comparison of the genome of the oral pathogen Treponema denticola
RT with other spirochete genomes.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:5646-5651(2004).
DR EMBL; AE017253; AAS12701.1; -.
DR TIGR; TDE2181; -.
DR InterPro; IPR008934; AcPase.VanPerase.
DR InterPro; IPR000326; Pesterase_PA_PTP.
DR Pfam; PF01569; PAP2; 1.
DR SMART; SM00014; acidPPC; 1.
KW Complete proteome.
SQ SEQUENCE 346 AA; 38328 MW; 7CC63ECA7E8AD165 CRC64;

Query Match 53.9%; Score 41; DB 2; Length 346;
Best Local Similarity 57.1%; Pred. No. 1.8e+02;
Matches 8; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 3 KRKAANAALLPVLL 16
Db 151 KLIKIAAAILPLLV 164

RESULT 30
Q89SZ7 PRELIMINARY; PRT; 403 AA.
AC Q89SZ7;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Bll12253 protein.
GN OrderedLocusNames=bll12253;
OS Bradyrhizobium japonicum.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Bradyrhizobium.
OX NCBI_TaxID=375;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=USDA110;
RX MEDLINE=22484998; PubMed=12597275;
RA Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiumi T.,
RA Sasamoto S., Watanabe A., Idesawa K., Iriguchi M., Kawashima K.,
RA Kohara M., Matsumoto M., Shimpo S., Tauruoka H., Wada T., Yamada M.,
RA Tabata S.;
RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium
RT Bradyrhizobium japonicum USDA110.";
RL DNA Res. 9:189-197(2002).
DR EMBL; AP005943; BAC47518.1; -.
DR InterPro; IPR001828; ANF_receptor.
DR Pfam; PF01094; ANF_receptor; 1.
KW Complete proteome.
SQ SEQUENCE 403 AA; 43679 MW; 7FDD58F6F6A980AA CRC64;

Query Match 53.9%; Score 41; DB 2; Length 403;
Best Local Similarity 64.3%; Pred. No. 2.1e+02;
Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 4 RKAAANAALLPVLLA 17
Db 8 RLAAACILLPMLAA 21

RESULT 31
Q8FTS8 PRELIMINARY; PRT; 492 AA.
AC Q8FTS8;
DT 01-WAR-2003 (TrEMBLrel. 23, Created)
DT 01-WAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Putative transport protein.
GN OrderedLocusNames=CEI483;
OS Corynebacterium efficiens.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacteriaceae; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxID=152794;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=YS-314;
RX MEDLINE=22723752; PubMed=12840036; DOI=10.1101/gr.1285603;
RA Nishio Y., Nakamura Y., Kawarabayasi Y., Usuda Y., Kimura E.,
RA Sugimoto S., Matsui K., Yamagishi A., Kikuchi H., Ikeo K.,
RA Gojobori T.;
RT "Comparative complete genome sequence analysis of the amino acid
RT replacements responsible for the thermostability of Corynebacterium
RT efficiens.";
RL Genome Res. 13:1572-1579(2003).
DR EMBL; AF005219; BAC18293.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR007114; MFS.

```

```

DR PROSITE; PSS0850; MFS; 1.
KW Complete proteome.
SQ SEQUENCE 492 AA; 50962 MW; 235E0CF3DA8EA1FF CRC64;

Query Match
Best Local Similarity 53.9%; Score 41; DB 2; Length 492;
Matches 7; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

QY 4 KKAARAAVLLPVLL 17
DB 8 KKAARAAVLLPVLL 21

RESULT 32
Q9BGZ6 PRELIMINARY; PRT; 524 AA.
AC Q9BGZ6;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Hypothetical protein.
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Frontal lobe left;
RA Osada N., Hida M., Kusuda J., Tanuma R., Iseki K., Hirai M., Terao K.,
RA Suzuki Y., Sugano S., Hashimoto K.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB05261; BAB21885.1; -
DR InterPro; IPR001395; Aldo/ket_red.
DR InterPro; IPR008940; Prenyl_trans.
DR InterPro; IPR001440; TPR.
DR Pfam; PF00515; TPR_1; 5.
DR SMART; SM00028; TPR; 7.
DR PROSITE; PSS0063; ALDKETO_REDUCTASE_3; UNKNOWN_1.
DR PROSITE; PSS0005; TPR; 5.
DR PROSITE; PSS0293; TPR_REGION; 1.
KW Hypothetical protein; Repeat; TPR repeat.
SQ SEQUENCE 524 AA; 59225 MW; 0E5E5A95EB0BB60 CRC64;

Query Match
Best Local Similarity 53.9%; Score 41; DB 2; Length 524;
Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 2 KKRKAAAVLLPVLL 16
DB 220 KKKLIAAVLLGILL 234

RESULT 33
Q25230 PRELIMINARY; PRT; 528 AA.
AC Q25230;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-JUN-2003 (TREMBLrel. 24, Last annotation update)
DE Hypothetical protein HP0486.
OS OrderedLocustNames=HP0486;
GN Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; Epsilonproteobacteria; Campylobacteriales;
OC Helicobacteraceae; Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=26695 / ATCC 700392;
RX MEDLINE=97394467; PubMed=9252185; DOI=10.1038/41483;
RA Tomb J.-F., White O., Kervatage A.R., Clayton R.A., Sutton G.G.,
RA Fleischmann R.D., Ketchum K.A., Klenk H.-P., Gill S.R.,
RA Dougherty B.A., Nelson K.E., Quackenbush J., Zhou L., Kirkness E.F.,

```

```

RA Peterson S.N., Loftus B.J., Richardson D.L., Dodson R.J., Khalak H.G.,
RA Glodek A., McKenney K., FitzGerald L.M., Lee N., Adams M.D.,
RA Hickey E.K., Berg D.E., Gocayne J.D., Uterback T.R., Peterson J.D.,
RA Kelley J.M., Cotton M.D., Weidman J.F., Fujii C., Bowman C.,
RA Wattey L., Wallin E., Hayes W.S., Borodovsky M., Karp P.D.,
RA Smith H.O., Fraser C.M., Venter J.C.;
RT "The complete genome sequence of the gastric pathogen Helicobacter
RT pylori."
RL Nature 388:539-547(1997).
DR EMBL; AE000563; AAD07554.1; -
DR PIR; F64580; F64580.
DR TIGR; HP0486; -
DR GO; GO:0009279; C:outer membrane (sensu Gram-negative Bacteria); IEA.
DR InterPro; IPR003678; HP_OMP_2.
DR Pfam; PF02521; HP_OMP_2; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 528 AA; 59416 MW; ED37A2D668123430 CRC64;

Query Match
Best Local Similarity 53.9%; Score 41; DB 2; Length 528;
Matches 11; Conservative 2; Mismatches 1; Indels 4; Gaps 1;

QY 2 KKRKAAAVL-----LPVL 15
DB 4 KKRKAAVALLKRFTLPVL 21

RESULT 34
Q65MC3 PRELIMINARY; PRT; 529 AA.
AC Q65MC3;
DT 25-OCT-2004 (TREMBLrel. 28, Created)
DT 25-OCT-2004 (TREMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TREMBLrel. 28, Last annotation update)
DE Malp.
GN Name=malP; ORFNames=BLI00857;
OS Bacillus licheniformis DSM 13.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=279010;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DSM 13;
RX PubMed=15383718;
RA Veith B., Herzberg C., Steckel S., Feesche J., Maurer K.H.,
RA Ehrenreich P., Baumeister S., Henne A., Liesegang H., Merkl R.,
RA Ehrenreich A., Gottschalk G.;
RT "The Complete Genome Sequence of Bacillus licheniformis DSM13, an
RT Organism with Great Industrial Potential."
RL J. Mol. Microbiol. Biotechnol. 7:204-211(2004).
DR EMBL; AE017333; AAU39791.1; -
SQ SEQUENCE 529 AA; 58301 MW; B05175B6038A0407 CRC64;

Query Match
Best Local Similarity 53.9%; Score 41; DB 2; Length 529;
Matches 7; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 1 KKRKAAVLLPVLL 16
DB 299 KKKKIVAGLLIPITL 314

RESULT 35
Q62XR6 PRELIMINARY; PRT; 530 AA.
AC Q62XR6;
DT 25-OCT-2004 (TREMBLrel. 28, Created)
DT 25-OCT-2004 (TREMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TREMBLrel. 28, Last annotation update)
DE Phosphotransferase system (PTS) maltose-specific enzyme IICB
DE component.
GN Name=malP; ORFNames=BL03019;
OS Bacillus licheniformis DSM 13.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.

```

```

OX NCBI_TaxID=279010;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 14580;
RA Rey M.W., Ramaiya P., Nelson B.A., Brody-Karpin S.D., Zaretsky E.J.,
RA Tang M., de Leon A.L., Xiang H., Gusti V., Clausen I.G., Olsen P.B.,
RA Rasmussen M.D., Andersen J.T., Jorgensen P.L., Larsen T.S.,
RA Sorokin A., Bolotin A., Lapidus A., Galleron N., Ehrlich S.D.,
RA Berka R.M.;
RT "Complete genome sequence of the industrial bacterium Bacillus
RT licheniformis and comparisons with closely related Bacillus species.";
RL Genome Biol. 5:R77-R77(2004).
DR EMBL: CP000002; AAU22442.1; -.
KW Transferase.
SQ SEQUENCE 530 AA; 58432 MW; B045829108BA0407 CRC64;

Query Match 53.9%; Score 41; DB 2; Length 530;
Best Local Similarity 43.8%; Pred. No. 2.7e+02;
Matches 7; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 1 KKKKRAAAVLLPVLL 16
DB 300 KKKKIVAGLLIPITL 315

RESULT 36
MM14_PIG
ID MM14_PIG STANDARD; PRT; 580 AA.
AC Q9XT90;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Matrix metalloproteinase-14 precursor (EC 3.4.24.80) (MMP-14)
DE (Membrane-type-1 matrix metalloproteinase 1) (MT1-MMP) (MT1MMP1)
DE (Membrane-type-1 matrix metalloproteinase) (MT1-MMP) (MT1MMP).
GN Name=MMP14;
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99095929; PubMed=9881602; DOI=10.1016/S0945-053X(98)90098-1;
RA Caron C., Xue J., Bartlett J.D.;
RT "Expression and localization of membrane type 1 matrix
RT metalloproteinase in tooth tissues.";
RL Matrix Biol. 17:501-511(1998).
CC -1- FUNCTION: Seems to specifically activate progelatinase A. May thus
CC trigger invasion by tumor cells by activating progelatinase A on
CC the tumor cell surface (By similarity). May play a role in the
CC biomineralization of enamel and dentin.
CC -1- CATALYTIC ACTIVITY: Endopeptidase activity. Activates
CC progelatinase A by cleavage of the propeptide at 37-Asn-|-Leu-38.
CC Other bonds hydrolyzed include 35-Gly-|-Ile-36 in the propeptide
CC of collagenase 3, and 341-Asn-|-Phe-342, 441-Asp-|-Leu-442 and
CC 354-Gln-|-Thr-355 in the aggrecan interglobular domain.
CC -1- COFACTOR: Binds 1 zinc ion per subunit, calcium (By similarity).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein (Potential).
CC -1- TISSUE SPECIFICITY: Highly expressed in developing tooth tissues.
CC -1- SIMILARITY: Belongs to the peptidase M10A family.
CC -1- SIMILARITY: Contains 1 hemopexin-like domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AF067419; AAD38324.1; -.
CC HSP; P50281; 1BQ0.
CC MEROPS; M10.014; -.

```

```

DR InterPro; IPR000585; Hemopexin.
DR InterPro; IPR001818; Pept_M10A_M12B.
DR InterPro; IPR006025; Pept_M_Zn_BS.
DR InterPro; IPR006026; PeptIdase_M.
DR Pfam; PF00045; Hemopexin; 4.
DR Pfam; PF00413; Peptidase_M10; 1.
DR Pfam; PF03933; Peptidase_M10_N; 1.
DR PRINTS; PR00138; MATRXIN.
DR SMART; SM00120; HK; 4.
DR SMART; SM00235; ZMC; 1.
DR PROSITE; PS00546; CYSTEINE_SWITCH; 1.
DR PROSITE; PS00024; HEMOPEXIN; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; 1.
KW Calcium; Hydrolase; Metalloprotease; Signal; Transmembrane; Zinc;
KW Zymogen.
FT SIGNAL 1 28 Potential.
FT PROPEP 29 109 Activation peptide.
FT CHAIN 110 580 Matrix metalloproteinase-14.
FT DOMAIN 110 539 Extracellular (Potential).
FT TRANSMEM 540 560 Potential.
FT DOMAIN 561 580 Cytoplasmic (Potential).
FT DOMAIN 314 509 Hemopexin-like.
FT SITE 51 91 Cysteine switch (Potential).
FT METAL 237 237 Zinc (catalytic) (By similarity).
FT ACT_SITE 238 238 By similarity.
FT METAL 241 241 Zinc (catalytic) (By similarity).
FT METAL 247 247 Zinc (catalytic) (By similarity).
FT DISULFID 317 506 By similarity.
SQ SEQUENCE 580 AA; 65934 MW; B782C3C569A96CAC CRC64;

Query Match 53.9%; Score 41; DB 1; Length 580;
Best Local Similarity 81.8%; Pred. No. 3e+02;
Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 6 AAAAVLLPVLL 16
DB 536 SAAAVVLPVLL 546

RESULT 37
MM14_HUMAN
ID MM14_HUMAN STANDARD; PRT; 582 AA.
AC P50281; Q92678;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Matrix metalloproteinase-14 precursor (EC 3.4.24.80) (MMP-14)
DE (Membrane-type matrix metalloproteinase 1) (MT1-MMP) (MT1MMP1)
DE (Membrane-type-1 matrix metalloproteinase) (MT1-MMP) (MT1MMP) (MMP-
DE X1).
GN Name=MMP14;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=94286011; PubMed=8015608; DOI=10.1038/370061a0;
RA Sato H., Takino T., Okada Y., Cao J., Shinagawa A., Yamamoto E.,
RA Seiki M.;
RT "A matrix metalloproteinase expressed on the surface of invasive
RT tumour cells.";
RL Nature 370:61-65(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=95237627; PubMed=7721107; DOI=10.1016/0378-1119(94)00637-8;
RA Takino T., Sato H., Yamamoto E., Seiki M.;
RT "Cloning of a human gene potentially encoding a novel matrix
RT metalloproteinase having a C-terminal transmembrane domain.";
RL Gene 155:293-298(1995).
RN [3]

```



RP SEQUENCE FROM N.A.  
RX MEDLINE=95224014; PubMed=7708715;  
RA Okada A., Belloq J.P., Rouyer N., Chenard M.P., Rio M.C., Chambon P.,  
RA Basset P.;  
RT "Membrane-type matrix metalloproteinase (MT-MMP) gene is expressed in  
RT stromal cells of human colon, breast, and head and neck carcinomas.";  
RL Proc. Natl. Acad. Sci. U.S.A. 92:2730-2734(1995).  
RN [4]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Lung;  
RX MEDLINE=95377289; PubMed=7649159;  
RA Will H., Hinzmann B.;  
RT "cDNA sequence and mRNA tissue distribution of a novel human matrix  
RT metalloproteinase with a potential transmembrane segment.";  
RL Eur. J. Biochem. 231:602-608(1995).  
RN [5]  
RP SEQUENCE FROM N.A.  
RA Luo G.X., Reisfeld R.A., Strongin A.V.;  
RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.  
RN [6]  
RP SEQUENCE FROM N.A.  
RA Lohi J.L., Westermarck J., Kaehaeri V.M., Keski-Oja J.;  
RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.  
RN [7]  
RP SEQUENCE OF 112-116.  
RX PubMed=880434;  
RA Sato H., Kinoshita T., Takino T., Nakayama K., Seiki M.;  
RT "Activation of a recombinant membrane type 1-matrix metalloproteinase  
RT (MT1-MMP) by furin and its interaction with tissue inhibitor of  
RT metalloproteinases (TIMP)-2.";  
RL FEBS Lett. 393:101-104(1996).  
CC -!- FUNCTION: Seems to specifically activate progelatinase A. May thus  
CC trigger invasion by tumor cells by activating progelatinase A on  
CC the tumor cell surface.  
CC -!- CATALYTIC ACTIVITY: Endopeptidase activity. Activates  
CC progelatinase A by cleavage of the propeptide at 37-Asn-Leu-38.  
CC Other bonds hydrolyzed include 35-Gly-Ile-36 in the propeptide  
CC of collagenase 3, and 341-Asn-Phe-342, 441-Asp-Leu-442 and  
CC 354-Gln-Thr-355 in the aggrecan interglobular domain.  
CC -!- COFACTOR: Binds 1 zinc ion per subunit, calcium (By similarity).  
CC -!- SUBCELLULAR LOCATION: Type I membrane protein (Potential).  
CC -!- TISSUE SPECIFICITY: In stromal cells of colon, breast, and head  
CC and neck.  
CC -!- SIMILARITY: Belongs to the peptidase M10A family.  
CC -!- SIMILARITY: Contains 1 hemopexin-like domain.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; D26512; BAA05519.1; -  
DR EMBL; X83535; CAA58519.1; -  
DR EMBL; 248481; CAA88372.1; -  
DR EMBL; U41078; AAB3770.1; -  
DR EMBL; X90925; CAA62432.1; -  
DR PIR; I38028; I38028.  
DR PDB; 1BQQ; X-ray; M=114-287.  
DR PDB; 1BUV; X-ray; M=114-287.  
DR MEROPS; M10.014; -.  
DR Genew; HGNC:7160; MMP14.  
DR MIM; 600754; -.  
DR GO; GO:0005887; C:integral to plasma membrane; TAS.  
DR GO; GO:0004222; F:metalloendopeptidase activity; TAS.  
DR GO; GO:0008270; F:zinc ion binding; TAS.  
DR GO; GO:0006508; P:proteolysis and peptidolysis; TAS.  
DR InterPro; IPR000585; Hemopexin.  
DR InterPro; IPR001818; Pept\_M10A\_M12B.  
DR InterPro; IPR006025; Pept\_M\_Zn\_BS.  
DR InterPro; IPR006026; Peptidase\_M.

DR Pfam; PF00045; Hemopexin; 4.  
DR Pfam; PF00413; Peptidase M10; 1.  
DR Pfam; PF03933; Peptidase M10\_N; 1.  
DR PRINTS; PR00138; MAVRIXIN.  
DR SMART; SM00120; HK; 4.  
DR SMART; SM00235; ZmC; 1.  
DR PROSITE; PS00546; CYSTEINE\_SWITCH; 1.  
DR PROSITE; PS00024; HEMOPEXIN; 1.  
DR PROSITE; PS00142; ZINC\_PROTEASE; 1.  
DR 3D-structure; Calcium; Direct protein sequencing; Hydrolase;  
KW Metalloprotease; Signal; Transmembrane; Zinc; Zymogen.  
FT SIGNAL 1 20  
FT PROPEP 21 111 Activation peptide.  
FT CHAIN 112 582 Matrix metalloproteinase-14.  
FT DOMAIN 112 541 Extracellular (Potential).  
FT TRANSMEM 542 562 Potential.  
FT DOMAIN 563 582 Cytoplasmic (Potential).  
FT DOMAIN 316 511 Hemopexin-like.  
FT SITE 93 93 Cysteine switch (Potential).  
FT METAL 239 239 Zinc (catalytic) (By similarity).  
FT ACT\_SITE 240 240 By similarity.  
FT METAL 243 243 Zinc (catalytic) (By similarity).  
FT METAL 249 249 Zinc (catalytic) (By similarity).  
FT DISULFID 319 508 By similarity.  
FT CONFLICT 338 338 K -> E (in Ref. 2, 4, 5 and 6).  
FT CONFLICT 500 500 S -> P (in Ref. 6).  
FT STRAND 123 125  
FT STRAND 127 128  
FT TURN 133 135  
FT HELIX 137 154  
FT STRAND 158 160  
FT TURN 169 170  
FT STRAND 178 182  
FT STRAND 198 202  
FT TURN 208 211  
FT TURN 213 216  
FT TURN 217 218  
FT STRAND 222 222  
FT TURN 224 225  
FT STRAND 232 232  
FT HELIX 233 243  
FT TURN 244 246  
FT TURN 253 254  
FT STRAND 260 260  
FT TURN 267 268  
FT HELIX 273 276  
FT TURN 277 278  
FT HELIX 279 282  
FT TURN 283 285  
SQ SEQUENCE 582 AA; 65983 MW; EFCEDCE6A41116P9 CRC64;

Query Match 53.9%; Score 41; DB 1; Length 582;  
Best Local Similarity 81.8%; Pred. No. 3e+02;  
Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 6 AAAAVLLPVL 16  
:|||||

Db 538 SAAAVLPLVL 548

RESULT 38

MM14 MOUSE  
ID MM14 MOUSE STANDARD; PRT; 582 AA.  
AC P53690; O08645; O35369;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 25-OCT-2004 (Rel. 45, Last annotation update)  
DE Matrix metalloproteinase-14 precursor (EC 3.4.24.80) (MMP-14)  
DE (Membrane-type matrix metalloproteinase 1) (MT-MMP 1) (MTMMP1)  
DE (Membrane-type-1 matrix metalloproteinase) (MT1-MMP) (MMP-X1)  
DE (MT-MMP).  
DE Name=Mmp14; Synonyms=Mtmmp;  
GN Mus musculus (Mouse).  
OS

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RN SEQUENCE FROM N.A.  
RP MEDLINE=95224014; PubMed=7708715;  
RX Okada A., Bellocq J.P., Rouyer N., Chenard M.P., Rio M.C., Chambon P.,  
RA Bassot P.;  
RA "Membrane-type matrix metalloproteinase (MT-MMP) gene is expressed in  
RT stromal cells of human colon, breast, and head and neck carcinomas.";  
RL Proc. Natl. Acad. Sci. U.S.A. 92:2730-2734(1995).  
RN [2]  
RN REVISIONS.  
RP Odaka A.;  
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RN SEQUENCE FROM N.A.  
RC STRAIN=129/Sv;  
RX MEDLINE=97467335; PubMed=9325265; DOI=10.1074/jbc.272.41.25511;  
RA Apte S.S., Fukui N., Beier D.R., Olsen B.R.;  
RT "The matrix metalloproteinase-14 (MMP-14) gene is structurally  
RT distinct from other MMP genes and is co-expressed with the TIMP-2 gene  
RT during mouse embryogenesis.";  
RL J. Biol. Chem. 272:25511-25517(1997).  
RN [4]  
RN SEQUENCE FROM N.A.  
RC STRAIN=CD-1; TISSUE=Kidney;  
RX MEDLINE=98311877; PubMed=9648071;  
RA Ota K., Stetler-Stevenson W.G., Yang Q., Kumar A., Wada J.,  
RA Kashihara N., Wallner E.I., Kanwar Y.S.;  
RT "Cloning of murine membrane-type-1-matrix metalloproteinase (MT-1-MMP)  
RT and its metalloproteinase-type-1-matrix metalloproteinase (MT-1-MMP)  
RT its inhibitor.";  
RL Kidney Int. 54:131-142(1998).  
RN [5]  
RN FUNCTION.  
RX MEDLINE=99449306; PubMed=10520996; DOI=10.1016/S0092-8674(00)80064-1;  
RA Holbeck K., Bianco P., Caterina J., Yamada S., Kromer M.,  
RA Kuznetsov S.A., Mankani M., Robey P.G., Poole A.R., Pidoux I.,  
RA Ward J.M., Birkedal-Hansen H.;  
RT "MT1-MMP-deficient mice develop dwarfism, osteopenia, arthritis, and  
RT connective tissue disease due to inadequate collagen turnover.";  
RL Cell 99:81-92(1999).  
RN [6]  
RN FUNCTION: Endopeptidase that degrades various components of the  
CC extracellular matrix, such as collagen. Activates progelatinase A.  
CC Essential for pericellular collagenolysis and modeling of skeletal  
CC and extracellular connective tissues during development.  
CC  
CC -1- CATALYTIC ACTIVITY: Endopeptidase activity. Activates  
CC progelatinase A by cleavage of the propeptide at 37-Asn-Leu-38.  
CC Other bonds hydrolyzed include 35-Gly-Ile-36 in the propeptide  
CC of collagenase 3, and 341-Asn-Ile-Phe-342, 441-Asp-Ileu-442 and  
CC 354-Gln-Thr-355 in the aggrecan interglobular domain.  
CC  
CC -1- COFACTOR: Binds 1 zinc ion per subunit, calcium (By similarity).  
CC  
CC -1- SUBCELLULAR LOCATION: Type I membrane protein (Potential).  
CC  
CC -1- TISSUE SPECIFICITY: Highly expressed in placenta, kidney, heart,  
CC lung, embryonic skeletal and periskeletal tissues.  
CC prominently expressed in large arteries and the umbilical  
CC arteries, expressed at lower levels in the myocardium,  
CC craniofacial mesenchyme, nasal epithelium and liver capsule. At  
CC days 14.5 and 17.5, expressed in the musculoskeletal system, and  
CC ossification areas, with continued expression in the arterial  
CC tunica media.  
CC  
CC -1- PTM: The precursor is cleaved by a furin endopeptidase (By  
CC similarity).  
CC  
CC -1- SIMILARITY: Belongs to the peptidase M10A family.  
CC  
CC -1- SIMILARITY: Contains 1 hemopexin-like domain.  
CC  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC DR EMBL; X83536; CA58520.2; -.  
CC DR EMBL; AF022432; AAB86602.1; -.  
CC DR EMBL; AF022424; AAB86602.1; JOINED.  
CC DR EMBL; AF022425; AAB86602.1; JOINED.  
CC DR EMBL; AF022426; AAB86602.1; JOINED.  
CC DR EMBL; AF022427; AAB86602.1; JOINED.  
CC DR EMBL; AF022428; AAB86602.1; JOINED.  
CC DR EMBL; AF022429; AAB86602.1; JOINED.  
CC DR EMBL; AF022430; AAB86602.1; JOINED.  
CC DR EMBL; AF022431; AAB86602.1; JOINED.  
CC DR EMBL; U54984; AAB51753.1; -.  
CC DR HSSP; P50281; 1BQQ.  
CC DR MEROPS; M10.014; -.  
CC DR MGD; MGI:101900; Mmp14.  
CC DR InterPro; IPR000585; Hemopexin.  
CC DR InterPro; IPR001818; Pept\_M10A\_M12B.  
CC DR InterPro; IPR006025; Pept\_M\_Zn\_BS.  
CC DR InterPro; IPR006026; Peptidase\_M.  
CC DR Pfam; PF00045; Hemopexin; 4.  
CC DR Pfam; PF00413; Peptidase M10; 1.  
CC DR Pfam; PF03933; Peptidase M10\_N; 1.  
CC DR PRINTS; PR00138; MATRIXIN.  
CC DR SMART; SM00120; HX; 4.  
CC DR SMART; SM00235; ZmC; 1.  
CC DR PROSITE; PS00546; CYSTEINE\_SWITCH; 1.  
CC DR PROSITE; PS00024; HEMOPEXIN; 1.  
CC DR PROSITE; PS00142; ZINC\_PROTEASE; 1.  
CC KW Calcium; Collagen degradation; Extracellular matrix; Hydrolase;  
KW Metalloprotease; Signal; Transmembrane; Zinc; Zymogen.  
FT SIGNAL 1 20  
FT PROPEP 21 111 By similarity.  
FT CHAIN 112 582 Matrix metalloproteinase-14.  
FT DOMAIN 112 541 Extracellular (Potential).  
FT TRANSMEM 542 562 Potential.  
FT DOMAIN 563 582 Cytoplasmic (Potential).  
FT SITE 93 93 Hemopexin-like.  
FT METAL 239 239 Cysteine switch (Potential).  
FT ACT\_SITE 240 240 Zinc (catalytic) (By similarity).  
FT METAL 243 243 Zinc (catalytic) (By similarity).  
FT METAL 249 249 By similarity.  
FT DISULFID 319 508 P -> S (in Ref. 3).  
FT CONFLICT 133 133 A -> D (in Ref. 1).  
FT CONFLICT 255 255 S -> A (in Ref. 3).  
FT CONFLICT 258 258 F -> L (in Ref. 1).  
FT CONFLICT 341 341 N -> P (in Ref. 3).  
FT CONFLICT 346 346 K -> T (in Ref. 1).  
FT CONFLICT 378 378 FD -> CV (in Ref. 1).  
FT CONFLICT 390 391 PK -> AN (in Ref. 1).  
FT CONFLICT 400 401 G -> V (in Ref. 1).  
FT CONFLICT 407 407 T -> S (in Ref. 1).  
FT CONFLICT 412 412 A -> T (in Ref. 1).  
FT CONFLICT 417 417 G -> R (in Ref. 4).  
FT CONFLICT 512 512  
SQ SEQUENCE 582 AA; 65935 MW; 3AB355158B4DD175 CRC64;  
Query Match 53.9%; Score 41; DB 1; Length 582;  
Best Local Similarity 81.8%; Pred. No. 3e+02;  
Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
QY 6 AAAAVLLPVLL 16  
Db 538 SAAAVVLPVLL 548  
RESULT 39  
MM14\_RABIT  
ID MM14\_RABIT STANDARD; PRT; 582 AA.  
AC Q95220; P79225;  
DT 01-NOV-1997 (Rel. 35, Created)

DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 25-OCT-2004 (Rel. 45, Last annotation update)  
 DE Matrix metalloproteinase-14 precursor (EC 3.4.24.80) (MMP-14)  
 DE (Membrane-type matrix metalloproteinase 1) (MT-MMP 1) (MTMMP1)  
 DE (Membrane-type-1 matrix metalloproteinase) (MT1-MMP) (MT1MMP).  
 GN Name=MMP14;  
 OS Oryctolagus cuniculus (Rabbit).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
 OX NCBI\_TaxID=9986;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=New Zealand white; TISSUE=Vascular smooth muscle;  
 RA Wang H., Keiser J.;  
 RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE OF 1-572 FROM N.A.  
 RC STRAIN=New Zealand white;  
 RA Sato T.;  
 RL Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: Seems to specifically activate progelatinase A. May thus  
 CC trigger invasion by tumor cells by activating progelatinase A on  
 CC the tumor cell surface (By similarity).  
 CC -1- CATALYTIC ACTIVITY: Endopeptidase activity. Activates  
 CC progelatinase A by cleavage of the propeptide at 37-Asn-Leu-38.  
 CC Other bonds hydrolyzed include 35-Gly-Ile-36 in the propeptide  
 CC of collagenase 3, and 341-Asn-Phe-342, 441-Asp-Leu-442 and  
 CC 354-Gln-Thr-355 in the aggrecan interglobular domain.  
 CC -1- COFACTOR: Binds 1 zinc ion per subunit, calcium (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein (Potential).  
 CC -1- SIMILARITY: Belongs to the peptidase M10A family.  
 CC -1- SIMILARITY: Contains 1 hemopexin-like domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; U83918; AAB41500.1; -;  
 DR EMBL; U73940; AAD13803.1; -;  
 DR HSSP; P50281; 1BQQ.  
 DR MEROPS: M10.014; -;  
 DR InterPro; IPR000585; Hemopexin.  
 DR InterPro; IPR001818; Pept\_M10A\_M12B.  
 DR InterPro; IPR006025; Pept\_M\_Zn\_BS.  
 DR InterPro; IPR006026; Peptidase M.  
 DR Pfam; PF00045; Hemopexin; 4.  
 DR Pfam; PF00413; Peptidase M10; 1.  
 DR Pfam; PF03933; Peptidase M10\_N; 1.  
 DR PRINTS; PR00138; MATRXIN.  
 DR SMART; SM00120; HX; 4.  
 DR SMART; SM00235; ZnMC; 1.  
 DR PROSITE; PS00546; CYSTEINE SWITCH; 1.  
 DR PROSITE; PS00024; HEMOPEXIN; 1.  
 DR PROSITE; PS00142; ZINC\_PROTEASE; 1.  
 KW Calcium; Hydrolase; Metalloprotease; Signal; Transmembrane; Zinc;  
 KW Zymogen.  
 FT SIGNAL 1 20 Potential.  
 FT PROPEP 21 111 Activation peptide.  
 FT CHAIN 112 582 Matrix metalloproteinase-14.  
 FT DOMAIN 112 541 Extracellular (Potential).  
 FT TRANSMEM 542 562 Potential.  
 FT DOMAIN 563 582 Cytoplasmic (Potential).  
 FT DOMAIN 316 511 Hemopexin-like.  
 FT SITE 93 93 Cysteine switch (Potential).  
 FT METAL 239 239 Zinc (catalytic) (By similarity).  
 FT ACT\_SITE 240 240 By similarity.  
 FT METAL 243 243 Zinc (catalytic) (By similarity).  
 FT METAL 249 249 Zinc (catalytic) (By similarity).  
 FT DISULFID 319 508 By similarity.

FT CONFLICT 29 29 Q -> K (in Ref. 2).  
 FT CONFLICT 268 268 K -> N (in Ref. 2).  
 FT CONFLICT 270 270 L -> V (in Ref. 2).  
 FT CONFLICT 275 275 E -> D (in Ref. 2).  
 FT CONFLICT 292 292 RCLLN -> KMPPP (in Ref. 2).  
 FT CONFLICT 298 300 QQP -> RTT (in Ref. 2).  
 FT CONFLICT 302 308 GLLPRIS -> RTLPDK (in Ref. 2).  
 FT CONFLICT 310 310 G -> R (in Ref. 2).  
 FT CONFLICT 317 317 K -> N (in Ref. 2).  
 FT CONFLICT 329 329 F -> L (in Ref. 2).  
 FT CONFLICT 360 360 L -> P (in Ref. 2).  
 SQ SEQUENCE 582 AA; 65963 MW; 844624B0AF1B6812 CRC64;  
 Query Match 53.9%; Score 41; DB 1; Length 582;  
 Best Local Similarity 81.8%; Pred. NO. 3e+02;  
 Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
 QY 6 AAAAVLLPVLL 16  
 Db 538 SAAAVVLPVLL 548  
 :||||:||||  
 RESULT 40  
 MM14\_RAT STANDARD; PRT; 582 AA.  
 ID MM14\_RAT  
 AC Q10739;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 25-OCT-2004 (Rel. 45, Last annotation update)  
 DE Matrix metalloproteinase-14 precursor (EC 3.4.24.80) (MMP-14)  
 DE (Membrane-type matrix metalloproteinase 1) (MT-MMP 1) (MTMMP1)  
 DE (Membrane-type-1 matrix metalloproteinase) (MT1-MMP) (MT1MMP) (MT-  
 DE MMP).  
 GN Name=Mmp14; Synonyms=Mtmmp;  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=95224014; PubMed=7708715;  
 RA Okada A., Bellocq J.P., Rouyer N., Chenard M.P., Rio M.C., Chambon P.,  
 RA Bassett P.;  
 RT "Membrane-type matrix metalloproteinase (MT-MMP) gene is expressed in  
 RT stromal cells of human colon, breast, and head and neck carcinomas.";   
 RL Proc. Natl. Acad. Sci. U.S.A. 92:2730-2734 (1995).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Cossins J., Clements J., Catlin G., Wells G.;  
 RL Submitted (SEP-1995) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: Seems to specifically activate progelatinase A. May thus  
 CC trigger invasion by tumor cells by activating progelatinase A on  
 CC the tumor cell surface.  
 CC -1- CATALYTIC ACTIVITY: Endopeptidase activity. Activates  
 CC progelatinase A by cleavage of the propeptide at 37-Asn-Leu-38.  
 CC Other bonds hydrolyzed include 35-Gly-Ile-36 in the propeptide  
 CC of collagenase 3, and 341-Asn-Phe-342, 441-Asp-Leu-442 and  
 CC 354-Gln-Thr-355 in the aggrecan interglobular domain.  
 CC -1- COFACTOR: Binds 1 zinc ion per subunit, calcium (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein (Potential).  
 CC -1- SIMILARITY: Belongs to the peptidase M10A family.  
 CC -1- SIMILARITY: Contains 1 hemopexin-like domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; X83537; CAA58521.1; -;  
 DR EMBL; X91785; CAA62897.1; -.

DR PIR; I84471; I84471.  
DR HSP; P50281; 1BQQ.  
DR MEROPS; M10.014; -.  
DR RGD; 620198; Mmp14.  
DR InterPro; IPR000585; Hemopexin.  
DR InterPro; IPR001818; Pept M10A M12B.  
DR InterPro; IPR006025; Pept M Zn BS.  
DR InterPro; IPR006026; Peptidase\_M.  
DR Pfam; PF00045; Hemopexin; 4.  
DR Pfam; PF00413; Peptidase M10; 1.  
DR Pfam; PF03933; Peptidase M10\_N; 1.  
DR PRINTS; PR00138; MATRINXIN.  
DR SMART; SM00235; ZnMc; 1.  
DR PROSITE; PS00546; CYSTEINE SWITCH; 1.  
DR PROSITE; PS00024; HEMOPEXIN; 1.  
DR PROSITE; PS00142; ZINC PROTEASE; 1.  
KW Calcium; Hydrolase; Metalloprotease; Signal; Transmembrane; Zinc;  
KW Zymogen.  
FT SIGNAL 1 20 Potential.  
FT PROPEP 21 111 Activation peptide.  
FT CHAIN 112 582 Matrix metalloproteinase-14.  
FT DOMAIN 112 541 Extracellular (Potential).  
FT TRANSMEM 542 562 Potential.  
FT DOMAIN 563 582 Cytoplasmic (Potential).  
FT DOMAIN 316 511 Hemopexin-like.  
FT SITE 93 93 Cysteine switch (Potential).  
FT METAL 239 239 Zinc (catalytic) (By similarity).  
FT ACT\_SITE 240 240 By similarity.  
FT METAL 243 243 Zinc (catalytic) (By similarity).  
FT METAL 249 249 Zinc (catalytic) (By similarity).  
FT DISULFID 319 508 By similarity.  
FT CONFLICT 68 68 I -> M (in Ref. 2).  
FT CONFLICT 255 255 D -> A (in Ref. 2).  
SQ SEQUENCE 582 AA; 66106 MW; 8B40FDD9999CA80C CRC64;

Query Match 53.9%; Score 41; DB 1; Length 582;  
Best Local Similarity 81.8%; Pred. No. 3e+02;  
Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 6 AAAAVLLPVLL 16  
:||||:|||||  
Db 538 SAAAVLPLVLL 548

RESULT 41  
Q9GLE4 PRELIMINARY; PRT; 582 AA.  
AC Q9GLE4;  
DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Membrane-type matrix metalloprotease 1.  
OS Bos taurus (Bovine).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
OC Bovinae; Bos.  
OX NCBI\_TaxID=9913;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Zhang B., Yan L., Moses M.A., Fang J., Miao H., Teang P.;  
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF290429; AAG28170.1; -.  
DR HSP; P50281; 1BQQ.  
DR MEROPS; M10.014; -.  
DR GO; GO:0005578; C:extracellular matrix (sensu Metazoa); IEA.  
DR GO; GO:0004222; F:metalloendopeptidase activity; IEA.  
DR GO; GO:0005508; P:proteolysis and peptidolysis; IEA.  
DR InterPro; IPR000585; Hemopexin.  
DR InterPro; IPR006026; Peptidase M.  
DR InterPro; IPR001818; Pept M10A M12B.  
DR InterPro; IPR006025; Pept M Zn BS.  
DR InterPro; IPR009070; PGBD\_like.

DR Pfam; PF00045; Hemopexin; 4.  
DR Pfam; PF03933; Peptidase M10\_N; 1.  
DR PRINTS; PR00138; MATRINXIN.  
DR SMART; SM00120; HX; 4.  
DR SMART; SM00235; ZnMc; 1.  
DR PROSITE; PS00546; CYSTEINE SWITCH; 1.  
DR PROSITE; PS00024; HEMOPEXIN; 1.  
DR PROSITE; PS00142; ZINC PROTEASE; 1.  
KW Metalloprotease; Protease.  
SQ SEQUENCE 582 AA; 65882 MW; 65174CE65D4040E1 CRC64;

Query Match 53.9%; Score 41; DB 2; Length 582;  
Best Local Similarity 81.8%; Pred. No. 3e+02;  
Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 6 AAAAVLLPVLL 16  
:||||:|||||  
Db 538 SAAAVLPLVLL 548

RESULT 42  
Q9XSP0 PRELIMINARY; PRT; 582 AA.  
AC Q9XSP0;  
DT 01-NOV-1999 (TrEMBLrel. 12, Created)  
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Membrane-type matrix metalloproteinase-1.  
GN Name=MtIMP;  
OS Capra hircus (Goat).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
OC Caprinae; Capra.  
OX NCBI\_TaxID=9925;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Placenta;  
RX MEDLINE=21593432; PubMed=11757910;  
RA Uekita T., Tanaka S., Sato H., Seiki M., Tojo H., Tachi C.;  
RT "Expression of membrane-type 1 matrix metalloproteinase (MT1-MMP) mRNA in trophoblast and endometrial epithelial cell populations of the synepitheliochorial placenta of goats (Capra hircus).";  
RL Arch. Histol. Cytol. 64:411-424(2001).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Placenta;  
RA Uekita T., Tanaka S., Mariko Y., Sato H., Seiki M., Tojo H., Tachi C.;  
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AB010921; BAA36551.1; -.  
DR HSP; P50281; 1BQQ.  
DR MEROPS; M10.014; -.  
DR GO; GO:0005578; C:extracellular matrix (sensu Metazoa); IEA.  
DR GO; GO:0004222; F:metalloendopeptidase activity; IEA.  
DR GO; GO:0005508; P:proteolysis and peptidolysis; IEA.  
DR InterPro; IPR000585; Hemopexin.  
DR InterPro; IPR006026; Peptidase M.  
DR InterPro; IPR001818; Pept M10A M12B.  
DR InterPro; IPR006025; Pept M Zn BS.  
DR InterPro; IPR009070; PGBD\_like.  
DR Pfam; PF00045; Hemopexin; 4.  
DR Pfam; PF03933; Peptidase M10\_N; 1.  
DR PRINTS; PR00138; MATRINXIN.  
DR SMART; SM00120; HX; 4.  
DR SMART; SM00235; ZnMc; 1.  
DR PROSITE; PS00546; CYSTEINE SWITCH; 1.  
DR PROSITE; PS00024; HEMOPEXIN; 1.  
DR PROSITE; PS00142; ZINC PROTEASE; 1.  
SQ SEQUENCE 582 AA; 65877 MW; B2B4E1CEFD7A4BE CRC64;

Query Match 53.9%; Score 41; DB 2; Length 582;  
Best Local Similarity 81.8%; Pred. No. 3e+02;  
Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 6 AAAAVALLPVLL 16  
:||||:|||||  
Db 538 SAAAVVLPVLL 548

## RESULT 43

Q6DFUS PRELIMINARY; PRT; 582 AA.  
AC Q6DFU5;  
DT 25-OCT-2004 (TRENBLrel. 28, Created)  
DT 25-OCT-2004 (TRENBLrel. 28, Last sequence update)  
DT 25-OCT-2004 (TRENBLrel. 28, Last annotation update)  
DE Matrix metalloproteinase 14 (Membrane-inserted).  
GN Name=Mmp14;  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6; TISSUE=Head;  
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Ussid T.B., Ioshizuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahy J., Helton E., Kettner M., Madan A., Rodrigues S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
RA Krzywinski M.I., Skalska U., Smalish D.E., Schnerch A., Schein J.E.,  
RA Jones S.J., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6; TISSUE=Head;  
RA Strausberg R.;  
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC076638; AAH76638.1; -;  
DR GO; GO:0005615; C:extracellular space; TAS.  
DR GO; GO:0016021; C:integral to membrane; TAS.  
DR InterPro; IPR000585; Hemopexin.  
DR InterPro; IPR006026; Peptidase M.  
DR InterPro; IPR001818; Pept M10A M12B.  
DR InterPro; IPR006025; Pept M10A M12B.  
DR Pfam; PF00045; Hemopexin; 4.  
DR Pfam; PF03933; Peptidase M10\_N; 1.  
DR PRINTS; PR00138; MATRXIN.  
DR SMART; SM00120; HK; 4.  
DR SMART; SM00235; ZnMc; 1.  
DR PROSITE; PS00546; CYSTEINE SWITCH; 1.  
DR PROSITE; PS00024; HEMOPEXIN; 1.  
DR PROSITE; PS00142; ZINC PROTEASE; UNKNOWN 1.  
SQ SEQUENCE 582 AA; 65931 MW; 28FCF84FC0ACD1B CRC64;

Query Match 53.9%; Score 41; DB 2; Length 582;  
Best Local Similarity 81.8%; Pred. No. 3e+02;  
Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 6 AAAAVALLPVLL 16  
:||||:|||||  
Db 538 SAAAVVLPVLL 548

RESULT 44  
Q8BTX2 PRELIMINARY; PRT; 582 AA.  
AC Q8BTX2;  
DT 01-MAR-2003 (TRENBLrel. 23, Created)  
DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)  
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)  
DE Mus musculus 2 days neonate thymus cells cDNA, RIKEN full-length enriched library, clone:E430018K19 product:matrix metalloproteinase 14 (membrane-inserted), full insert sequence.  
GN Name=Mmp14;  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=NOD; TISSUE=Thymus;  
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;  
RA Carninci P., Hayashizaki Y.;  
RT "High-efficiency full-length cDNA cloning.";  
RL Meth. Enzymol. 303:19-44(1999).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=NOD; TISSUE=Thymus;  
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;  
RA RIKEN FANTOM Consortium;  
RT "Functional annotation of a full-length mouse cDNA collection.";  
RL Nature 409:685-690(2001).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=NOD; TISSUE=Thymus;  
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;  
RA RIKEN FANTOM Consortium;  
RT "Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs.";  
RL Nature 420:563-573(2002).  
RN [4]  
RP SEQUENCE FROM N.A.  
RC STRAIN=NOD; TISSUE=Thymus;  
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;  
RA Carninci P., Shibata Y., Hayatsu M., Sugahara Y., Shibata K., Itoh M.,  
RA Konno H., Okazaki Y., Muramatsu N., Hayashizaki Y.;  
RT "Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes.";  
RL Genome Res. 10:1617-1630(2000).  
RN [5]  
RP SEQUENCE FROM N.A.  
RC STRAIN=NOD; TISSUE=Thymus;  
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;  
RA Shibata K., Itoh M., Aizawa K., Nagao K., Sasaki N., Carninci P.,  
RA Konno H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,  
RA Suni N., Ishii Y., Nakamura S., Hazama M., Nishino T., Harada A.,  
RA Yanamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,  
RA Fujiwaki S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,  
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,  
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;  
RT "RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer.";  
RL Genome Res. 10:1757-1771(2000).  
RN [6]  
RP SEQUENCE FROM N.A.  
RC STRAIN=NOD; TISSUE=Thymus;  
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;  
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,  
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,  
RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,  
RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,  
RA Katoh H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,  
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,  
RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,  
RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,  
RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,

```
RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
RA Tomaru A., Tova T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK084476; BAC40377.1; -.
DR HSSP; P50281; 1BQO.
DR DR MGD; MGI:101900; Mmp14.
DR GO; GO:0005615; C:extracellular space; TAS.
DR GO; GO:0016021; C:integral to membrane; TAS.
DR InterPro; IPR000585; Hemopexin.
DR InterPro; IPR006026; Peptidase_M.
DR InterPro; IPR001818; Pept_M10A_M12B.
DR InterPro; IPR006025; Pept_M_Zn_BS.
DR InterPro; IPR009070; PGED_like.
DR Pfam; PF00045; Hemopexin_4.
DR Pfam; PF03933; Peptidase_M10_N; 1.
DR PRINTS; PR00138; MATRIXIN.
DR SMART; SM00120; HK; 4.
DR SMART; SM00235; ZnMc; 1.
DR PROSITE; PS00546; CYSTEINE_SWITCH; 1.
DR PROSITE; PS00024; HEMOPEXIN; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN_1.
SQ SEQUENCE 582 AA; 65919 MW; 6CA95CFD5811B093 CRC64;

Query Match 53.9%; Score 41; DB 2; Length 582;
Best Local Similarity 81.8%; Pred. No. 3e+02;
Matches 9; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 6 AAAAVLLPVLL 16
Db 538 SAAAVLLPVLL 548
:||||:|||||

RESULT 45
Q6EQK1 PRELIMINARY; PRT; 886 AA.
AC Q6EQK1;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Serine/threonine-specific receptor protein kinase-like.
GN Names=OSJNBa0042H24.55-1; Synonyms=OSJNBa0042H24.55-1;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzoae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SOURCE FROM N.A.
RA Sasaki T., Matsumoto T., Katayose Y.;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
SQ SEQUENCE FROM N.A.
RA Sasaki T., Matsumoto T., Fujisawa M.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 9, fosmid
clone:OSJNBa0042H24.55-1;".
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: Belongs to the Ser/Thr protein kinase family.
DR EMBL; AP007202; BAD46707.1; -.
DR EMBL; AP007202; BAD46707.1; -.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004574; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR011009; Kinase_like.
DR InterPro; IPR011611; LRR.
DR InterPro; IPR007090; LRR_plant.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR InterPro; IPR008271; Ser_thr_pkin_AS.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00560; LRR_1; 3.
DR Pfam; PF00069; Pkinase; 1.
```

```
DR PRINTS; PR00019; LEURICHRPT.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR SMART; SM00219; TyrKc; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW ATP-binding; Kinase; Receptor; Serine/threonine-protein kinase;
KW Transferase.
SQ SEQUENCE 886 AA; 98476 MW; 51C734FA9038FF8A CRC64;

Query Match 53.9%; Score 41; DB 2; Length 886;
Best Local Similarity 55.6%; Pred. No. 4.4e+02;
Matches 10; Conservative 4; Mismatches 2; Indels 2; Gaps 1;

QY 1 KKKRKA-AAAVLLPVLL 16
Db 513 KKKRKAVALAVVVPVVI 530
|||||:|||||

RESULT 46
Q7W3J1 PRELIMINARY; PRT; 177 AA.
AC Q7W3J1;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein.
GN OrderedLocNames=BPP4042;
OS Bordetella parapertussis.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Alcaligenaceae; Bordetella.
OX NCBI_TaxID=519;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=12822 / ATCC BAA-587;
RX MEDLINE=22827954; PubMed=12910271; DOI=10.1038/ng1227;
RA Parkhill J., Sebaihia M., Preston A., Murphy L.D., Thomson N.R.,
RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,
RA Cerdano-Tarraga A.-M., Temple L., James K.D., Harris B., Quail M.A.,
RA Achtman M., Atkin R., Baker S., Basham D., Bason N., Cherevach I.,
RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,
RA Fellwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,
RA Leather S., Moule S., Norberczak H., O'Neil S., Ormond D., Price C.,
RA Rabinovitch E., Rutter S., Sanders M., Saunders D., Seeger K.,
RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,
RA Unwin L., Whitehead S., Barrall B.G., Maskell D.J.;
RT "Comparative analysis of the genome sequences of Bordetella pertussis,
RT Bordetella parapertussis and Bordetella bronchiseptica.";
RL Nat. Genet. 35:32-40(2003).
DR EMBL; BX640435; CAE39325.1; -.
DR InterPro; IPR003509; UPF0102.
DR Pfam; PF02021; UPF0102; 1.
DR TIGRFAMs; TIGR00252; UPF0102; 1.
KW Complete proteome.
SQ SEQUENCE 177 AA; 19268 MW; 2A1BDFD4964FFBBB CRC64;

Query Match 53.3%; Score 40.5; DB 2; Length 177;
Best Local Similarity 70.6%; Pred. No. 1.2e+02;
Matches 12; Conservative 2; Mismatches 2; Indels 1; Gaps 1;

QY 1 KKKRKA-AAAVLLPVLL 16
Db 109 KKKRLARAANLLPVLL 125
||:|:|||||

RESULT 47
Q7W6M6 PRELIMINARY; PRT; 177 AA.
ID Q7W6M6;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein.
```

GN OrderedLocusNames=BB4515;  
 OS Bordetella bronchiseptica (Alcaligenes bronchisepticus).  
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;  
 OC Alcaligenaceae; Bordetella.  
 NCBI\_TaxID=518;  
 RN [1]\_TaxID=518;  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=RB50 / ATCC BAA-588;  
 RX MEDLINE=22827954; PubMed=12910271; DOI=10.1038/ng1227;  
 RA Parkhill J., Sebahia M., Preston A., Murphy L.D., Thomson N.R.,  
 RA Harris D.E., Holden M.T.G., Churcher C.M., Bentley S.D., Mungall K.L.,  
 RA Cardeno-Tarraga A.-M., Temple L., James K.D., Harris B., Quail M.A.,  
 RA Achman M., Akin R., Baker S., Basham D., Bason N., Cherevach I.,  
 RA Chillingworth T., Collins M., Cronin A., Davis P., Doggett J.,  
 RA Fellwell T., Goble A., Hamlin N., Hauser H., Holroyd S., Jagels K.,  
 RA Leather S., Moule S., Norberczak H., O'Neil S., Ormond D., Price C.,  
 RA Rabinowitz E., Rutter S., Sanders M., Saunders D., Seeger K.,  
 RA Sharp S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K.,  
 RA Unwin L., Whitehead S., Barrall B.G., Maskell D.J.;  
 RT "Comparative analysis of the genome sequences of Bordetella pertussis,  
 RT Bordetella parapertussis and Bordetella bronchiseptica.";  
 RL Nat. Genet. 35:32-40(2003).  
 DR EMBL: BX640450; CAE34878.1; --  
 DR InterPro: IPR003509; UPF0102.  
 DR Pfam: PF02021; UPF0102; 1.  
 DR TIGRFAMs: TIGR00252; UPF0102; 1.  
 KW Complete proteome.  
 SQ SEQUENCE 177 AA; 19418 MW; BF6CD86DE3838146 CRC64;  
  
 Query Match 53.3%; Score 40.5; DB 2; Length 177;  
 Best Local Similarity 70.6%; Pred. No. 1.2e+02;  
 Matches 12; Conservative 2; Mismatches 2; Indels 1; Gaps 1;  
  
 QY 1 KKKRKA-AAVLLPVLL 16  
 DB 109 KQGLRARAALLPVLL 125  
  
 RESULT 48  
 Q9VNM7 PRELIMINARY; PRT; 288 AA.  
 AC Q9VNM7;  
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
 DE 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)  
 DE CG1153-PA.  
 GN ORFNames=CG1153;  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.H., Blazer R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,  
 RA Abil J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borkova D., Botchan M.R., Bouck J., Brocktein P., Brotter P.,  
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,

RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,  
 RA Jatali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milhina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nussekern D.R., Pacleb J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,  
 RA Williams N.M., Woodage, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,  
 RA Yeh R.P., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of Drosophila melanogaster.";  
 RL Science 287:2185-2195(2000).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22426065; PubMed=12537568;  
 RA Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,  
 RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,  
 RA George R.A., Hoskins R.A., Lavery T., Muzny D.M., Nelson C.R.,  
 RA Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,  
 RA Svirskaas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,  
 RA Weinstock G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;  
 RT "Finishing a whole-genome shotgun: Release 3 of the Drosophila  
 RT melanogaster euchromatic genome sequence.";  
 RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22426070; PubMed=12537573;  
 RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J., Svirskaas R.,  
 RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,  
 RA Ashburner M., Celniker S.E.;  
 RT "The transposable elements of the Drosophila melanogaster euchromatin:  
 RT a genomics perspective.";  
 RL Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22426069; PubMed=12537572;  
 RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,  
 RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochownik S.E.,  
 RA Smith C.D., Tupy J.L., Whitfield B.J., Bayraktaroglu L., Berman B.P.,  
 RA Bettencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,  
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,  
 RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
 RA Lewis S.E.;  
 RT "Annotation of the Drosophila melanogaster euchromatic genome: a  
 RT systematic review.";  
 RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RG FlyBase;  
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RG FlyBase;  
 RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AE003600; AAF51902.1; --  
 DR FlyBase: FBgn0037414; Osi7.  
 SQ SEQUENCE 288 AA; 30036 MW; 80F7727D4EE61EAL CRC64;  
  
 Query Match 53.3%; Score 40.5; DB 2; Length 288;  
 Best Local Similarity 42.9%; Pred. No. 1.9e+02;  
 Matches 12; Conservative 2; Mismatches 1; Indels 13; Gaps 1;  
  
 QY 2 KKKRKA-----AAVLLPVLL 16  
 DB 162 KKKRKAATILGPILALVALKAAALLPVLL 189

```

RESULT 49
Q8JND4
ID Q8JND4 PRELIMINARY; PRT; 497 AA.
AC Q8JND4;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE ORF1864 (Fragment).
GN Name=orf1864;
OS White spot syndrome virus (WSSV).
OC Viruses; dsDNA viruses, no RNA stage; Nimaviridae; Whispovirus.
OX NCBI_TaxID=92652;
RN [1]
RP SEQUENCE FROM N.A.
RA Zhang X., Huang C., Hew C.L.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY044842; AAL02263.1; -.
DR InterPro; IPR000425; MIP.
DR PROSITE; PS00221; MIP; UNKNOWN_1.
DR NON_TER 497 497
SQ SEQUENCE 497 AA; 56616 MW; E8E5601E2FBB7DB3 CRC64;

Query Match 53.3%; Score 40.5; DB 2; Length 497;
Best Local Similarity 66.7%; Pred.No. 3.1e+02;
Matches 10; Conservative 2; Mismatches 2; Indels 1; Gaps 1;

QY 1 KKRRKAAAA-VLLPV 14
|:| | | | |
Db 318 KRKNTAAANILLPV 332

RESULT 50
Q8QOY9
ID Q8QOY9 PRELIMINARY; PRT; 544 AA.
AC Q8QOY9;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE VP544 (Fragment).
OS White spot syndrome virus (WSSV).
OC Viruses; dsDNA viruses, no RNA stage; Nimaviridae; Whispovirus.
OX NCBI_TaxID=92652;
RN [1]
RP SEQUENCE FROM N.A.
RA Zhang X., Huang C., Hew C.L.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF493145; AAM12816.1; -.
DR InterPro; IPR000425; MIP.
DR PROSITE; PS00221; MIP; UNKNOWN_1.
DR NON_TER 544 544
SQ SEQUENCE 544 AA; 61771 MW; CD1881F5BE294E02 CRC64;

Query Match 53.3%; Score 40.5; DB 2; Length 544;
Best Local Similarity 66.7%; Pred.No. 3.4e+02;
Matches 10; Conservative 2; Mismatches 2; Indels 1; Gaps 1;

QY 1 KKRRKAAAA-VLLPV 14
|:| | | | |
Db 365 KRKNTAAANILLPV 379

Search completed: June 2, 2005, 01:40:04
Job time : 135.594 secs

```